THE *GENIUS LOCI*[1] OF CRIME:
REVEALING ASSOCIATIONS IN TIME AND SPACE

by

Jeremy H. Ratcliffe B.Sc. FRGS

A thesis submitted to the University of Nottingham

for the degree of

Doctor of Philosophy

March, 1999

[1] The presiding deity of a place; hence the spirit or associations predominant in a locality, community or institution. *Concise English Dictionary, p. 484.*

Singularity is almost invariably a clue. The more featureless and commonplace a crime is, the more difficult is it to bring it home.

Sir Arthur Conan Doyle, *The Adventures of Sherlock Holmes*

UNIVERSITY OF NOTTINGHAM

Abstract

THE *GENIUS LOCI* OF CRIME:
REVEALING ASSOCIATIONS IN TIME AND SPACE

by Jeremy H. Ratcliffe B.Sc. FRGS

In most police services the only spatial and temporal analysis of crime was conducted until recently by statisticians at the force headquarters, with little or no regard for any short term or localised patterns of crime. In recent years there has been a move towards a more decentralised, proactive style of British policing focused at the police divisional and community level. This has left an intelligence void where force level analysis techniques are neither appropriate nor subtle enough to elicit any meaningful information at a local level from the mass of crime data generated within the police service.

This thesis reveals patterns in community level crime which have not been recognised previously using traditional techniques in spatial and temporal investigation which tend to lack the necessary analytical ability. Current policing considerations are recognised and the thesis concentrates on three aspects of police crime concern: accurate temporal analysis, repeat victimisation, and the identification of hotspots.

A number of new techniques are presented which are designed with the needs of a crime analyst at a divisional police station in mind, an individual who has until now lacked the necessary analytical tools to perform the role effectively.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PUBLICATIONS AND PRESENTATIONS FROM THIS THESIS

Ratcliffe, J.H. and McCullagh, M.J. (1998), **Aoristic crime analysis,** *International Journal of Geographical Information Science*, 12 (7): 751-764.

Ratcliffe, J.H. and McCullagh, M.J. (1998), **Identifying repeat victimisation with GIS,** *British Journal of Criminology*, 38 (4): 651-662.

Ratcliffe, J.H. and McCullagh, M.J. (1998), **The perception of crime hot spots: A spatial study in Nottingham, UK,** in N. LaVigne and J. Wartell, eds., *Crime Mapping Case Studies: Successes in the Field.* Police Executive Research Forum: Washington DC: 45-51.

Ratcliffe, J.H. (1998) **High volume crime and hotspots: Evaluating the intelligence dissemination process,** paper presented to 2nd Annual Crime Mapping Conference '*Mapping Out Crime*', National Institute of Justice, 12th December 1998, Washington D.C.

Ratcliffe, J.H. and McCullagh, M.J. (1999), **Burglary, Victimisation and Social Deprivation,** *Crime Prevention and Community Safety*, 1 (2): 37-45.

Ratcliffe, J.H. and McCullagh, M.J. (under review), **Hotbeds of crime and the search for spatial accuracy,** *Geographical Analysis.*

CONTENTS OF THE ACCOMPANYING CD-ROM

The accompanying CD-ROM contains two directories and a number of files in the root system.

The **Aaplay** directory contains a freeware animation player called Autodesk Animation Player for Windows, version 1.1. It may be required to view the animation files which have a *.FLC extension. The **Spam** directory contains two versions of the SPAM program referred to in Chapters 8 and 9, as well as two sample data text files.

The root directory of the CD-ROM contains two *.FLC animations, an AVI animation and the compiled code of a repeat location finding MapBasic program. These files are referred to in the text of the thesis.

For detailed instructions on installing and running this software, see the README.TXT file in the root directory of the CD-ROM.

ACKNOWLEDGEMENTS

The author wishes to thank a number of people without whose assistance this thesis would not have reached completion. Chief Superintendent Eddie Curtis and Assistant Chief Constable (Designate) Dr. Tom Williamson, both of Nottinghamshire Constabulary, were both helpful and positive towards the whole project. To maintain their commitment and willingness in the face of considerable obstructions from within their own force is a mark of their dedication to the whole business of fighting crime.

My gratitude also to Dr. Michael McCullagh for his enthusiasm and endless patience. His tolerance of my numerous blank looks and willingness to include me in the many aspects of his academic life have given me a valuable insight into academia and this thesis is the better for it. My thanks go to a supervisor (needs more work...), and a friend.

Finally I would like to dedicate this work to my wife, Philippa, who tolerated the whole nonsense with considerable patience and good humour.

GLOSSARY

British Crime Survey	Biannual survey of approximately 16,500 households in the UK to attain a clearer picture of the true level of crime in the country.
Comparative Case Analysis	Seeks to detect similarities between crimes that point to the same perpetrator or groups of perpetrators.
Crime Pattern Analysis (CPA)	Attempts to form a picture of the nature and scale of crime in a particular area. The analysis of crime distribution for significant patterns. This term generally refers to spatial patterns, though is often used to indicate mapping and not analysis.
Dark figure	The amount of unreported and undiscovered crime that does not make it into police crime records.
Environmental Systems Research Institute (ESRI)	Institute responsible for the creation and marketing of the ArcInfo and ArcView suite of GIS programs.
Geographical Information Science (GIS)	The science associated with the use of Geographical Information Systems.
Geographical Information System (GIS)	A computer program specifically designed to input, store, manipulate, analyse and display geographical data.
Hotspot	A significant concentration of crime incidents, focussed in a cluster where crime is higher than in other areas.
Hotspot analysis	The search for the highest concentration of hotspots.
Local Crime Analysis	The application of Comparative Case Analysis and Crime Pattern Analysis at a local level.
Modifiable Areal Unit Problem (MAUP)	How choice of aggregation technique and display resolution can affect the eventual display and interpretation of mapped data.
Modus Operandi	The method used to commit a crime.
Postcode Address File (PAF)	A digital file of households and their postcodes. The sample frame used for the British Crime Surveys.
Spatial and Temporal Analysis of Crime (STAC)	Spatial and Temporal Analysis of Crime. DOS based software developed by the Chicago justice department.

Spatial Crime Analysis System (SCAS)	Spatial Crime Analysis System. Windows-based crime analysis software which requires ArcView 3.0 to run. Developed by Crime Mapping Research Center, US Department of Justice.
Spatial Pattern Analysis Machine (SPAM)	Windows based software developed for this thesis to detect local statistically significant crime hotspots.
Travel to crime	Travel to crime studies tend to focus on the distance a criminal is prepared to travel to perpetrate a criminal offence and rely on arrest records of individuals or occasionally interviews of prisoners
Vertical Mapper	Add-on software for MapInfo which permits the analysis and display of irregularly spaced surface data. Mainly used for limited surface modelling.

ABBREVIATIONS

ABH	Actual Bodily Harm
ACF	Autocorrelation function
ACPO	Association of Chief Police Officers
ArcInfo	Commercial workstation orientated GIS package
ArcView	Commercial PC-based GIS package (derived from ArcInfo)
BCS	British Crime Survey
BTP	British Transport Police
CAD	Computer Aided Despatch
CBD	Central Business District
CPA	Crime Pattern Analysis
CRIS	Crime Recording Interim/Information System
ESRC	Economic and Social Research Council
ESRI	Environmental Systems Research Institute
GAM	Geographical Analysis Machine
GIS	Geographical Information System/Science
ILC	Index of Local Conditions
LISA	Local Indicators of Spatial Association
MapInfo	Commercial GIS package.
MARS	Merseyside Address Referencing System
MAUP	Modifiable Areal Unit Problem
MBR	Minimum Bounding Rectangle
MFC	Microsoft Foundation Classes
MIS	Management/Merseyside Information System
MOE	Method of Entry (usually relates to burglaries)
OS	Ordnance Survey

OSAPR	Ordnance Survey Address-Point Reference
PAF	Postcode Address File
POE	Point of Entry (usually relates to burglaries)
SCAS	Spatial Crime Analysis System
SPAM	Spatial Pattern Analysis Machine
SPSS	Statistical Package for the Social Sciences (statistical software)
SQL	Standard Query Language
STAC	Spatial and Temporal Analysis of Crime
TGIS	Temporal GIS
TSA	Time Series Analysis
UPRN	Unique Property Reference Number

1. Introduction

This chapter describes the aims of the thesis and outlines the framework of the chapters to follow. Most crime analysis has taken place until recently at a force level, only investigating the long term trends in crime at the macro (county) level. A spatial crime analysis model is developed which investigates the required inputs and possible outputs of an analytical system which could be developed to operate in a local (divisional) police environment. The structure of the thesis is then described.

1.1. INTRODUCTION

This thesis examines the potential for using information technology (IT) and particularly geographical information science (GIS) in the analysis of local crime patterns. Analysis of crime has been conducted for many years in the UK, though mostly at a county level by statisticians at police force headquarters. In recent years there has been a move towards a more decentralised, problem-orientated style of policing in Britain and this has placed more focus on divisional crime analysts and intelligence officers. As the demands on the police service expand, with no significant increase in officer numbers, it has become essential that the police extract the maximum intelligence from the available data in order to have an impact on local crime by targeting local resources effectively.

Over recent years there has been a significant increase in the use of IT within the police service. Most forces in the UK that record crime information on computer now also log their incident data digitally. Crime reports are completed for any incidents that come to the attention of the police where an act of criminal legislation has been violated and the suspect could be indicted under criminal law. This excludes lesser crimes such as minor motor vehicle offences from being called 'crimes'. Incidents (also known as 'calls for service' in the US and Australia) are a record of requests for assistance from the public or from police officers, irrespective of whether the incident results in a crime report. In the UK this often happens when a person calls the police station either directly or through the '999' system. An incident will generally be created whenever any police activity is required and the incident log is a record of the day-to-day policing activity of an area. The knowledge of where incidents are happening is vital and a number of UK police forces have integrated GIS packages into their incident systems. Forces which have added GIS technology to their computerised Command and Control systems for incident handling include: the Metropolitan Police in London (Ireland, 1998), Nottinghamshire (Nottinghamshire Constabulary, 1997; ICL, 1995), Cleveland (Clegg and Robson, 1995), West Midlands (Page, 1997) and Hampshire (Fox-Clinch, 1997).

Most police forces in the UK now attempt to analyse some of their available data, and there are a number of packages (Harlequin, 1997; i2, 1997) which perform basic analysis (these are reviewed in the next chapter) though 'government funding for applied research has been insignificant compared to that in the US' (Ireland, 1998 p.21). This expansion in IT use resulted in a growth in the number of crime analysts employed within police forces and they can now sometimes be found not just in force headquarters but also in divisional stations. These individuals (either police officers or civilians) are responsible for the extraction of useful intelligence from the available data sources. The range of digital material available for analysis has increased rapidly in recent years with the inception of computer encoding of what were previously paper crime and incident records. The georeferencing of this crime and incident data has allowed the introduction of GIS as a tool for examining the spatial dimensions of crime.

There is a geography to crime and it is recognised that the distribution of offending and the demand for police service are not uniform, but that the 'ecology of crime is for the most part a very localized affair, a patchwork of 'hot spots' and quiet zones' (Morgan and Newburn, 1997, p.150). Knowledge of these crime patterns and their distribution is more important than ever to the police as the expanding role of the beat officers means that less time is available for routine patrolling. It has been estimated that on the basis of current police strength nationally, each patrolling police officer covers an area with an average of:

- 18,000 inhabitants
- 7,500 houses
- 23 pubs
- 9 schools
- 140 miles of pavement
- 77 miles of roads
- Over 100 prolific offenders

Source (Audit Commission, 1996)

With little or no formalised spatial training and few crime-specific analytical routines available to crime analysts the implementation of GIS within police forces has had varying degrees of success (Maltz *et al.*, 1991; Page, 1997; Read and Oldfield, 1995), and notable cases of failure (Openshaw *et al.*, 1990). Even with the lack of unqualified success within most forces, it is possible to construct a model which shows the current and potential benefits that GIS can bring to the crime analysis environment.

1.1.1. The spatial crime analysis model

Figure 1-1 shows a theoretical model of the potential impact that spatial crime analysis could have in a police environment. There are two main paths: low volume serial crime analysis (shown through the *specialist analyst* in Figure 1-1), and high volume crime analysis (shown through the *crime analyst* in Figure 1-1).

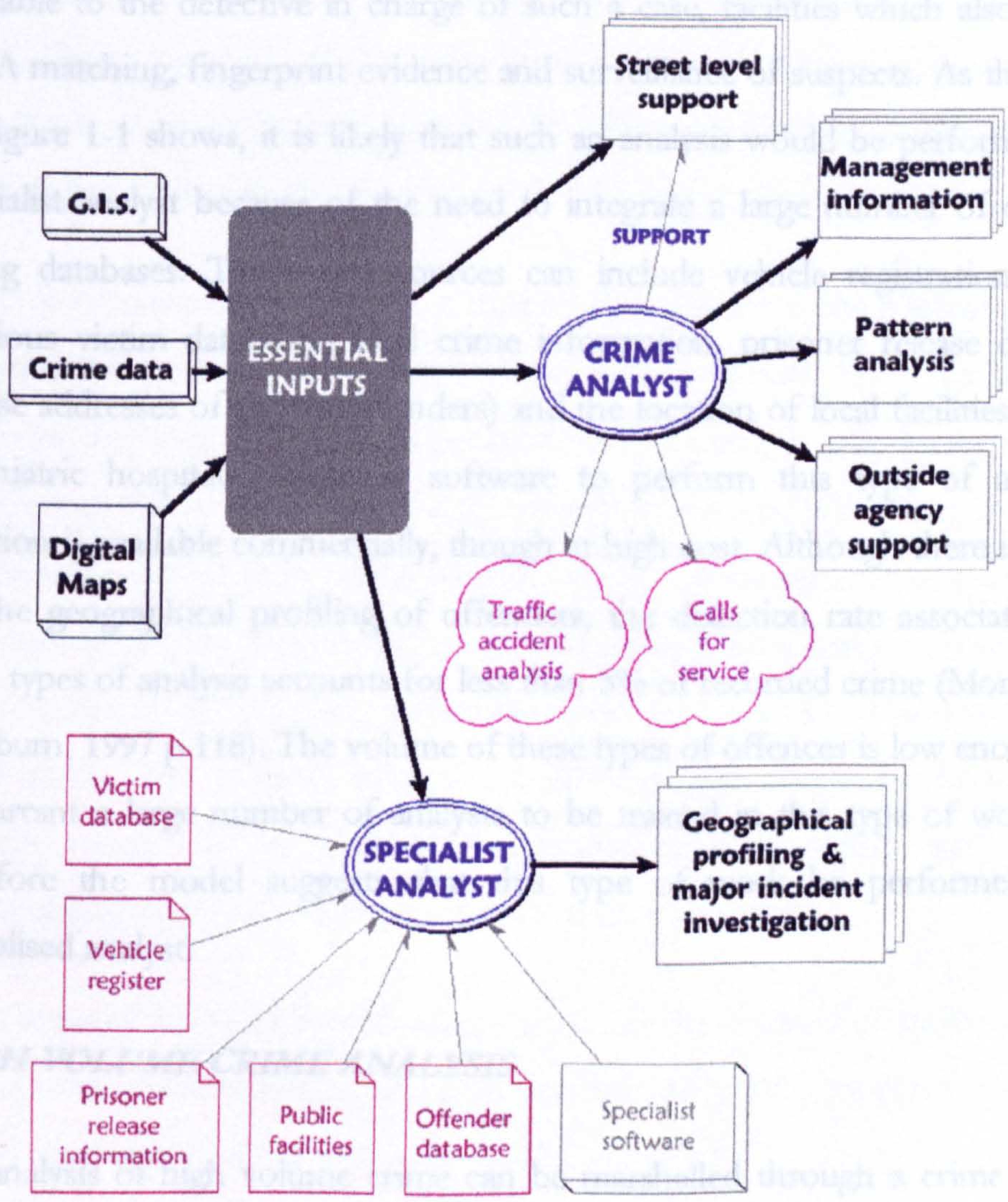


Figure 1-1 Spatial crime analysis model within a police environment.

The model in Figure 1-1 shows that there are three essential inputs to a spatial crime analysis system: a GIS, crime data, and digital maps. From this basis it is possible to follow either the low volume serial crime route or the high volume crime path.

Low volume serial crime analysis

Low volume serial crime analysis is an area that deals with the geographical profiling of offenders of serious crimes. Serial rapists and people who commit sexual offences against adults and children are rare in any society. However, the commission of any crimes such as these generates an understandably high degree of public concern and considerable effort on the part of the police to catch the persons responsible. Geographical offender profiling is one of many facilities available to the detective in charge of such a case, facilities which also include DNA matching, fingerprint evidence and surveillance of suspects. As the model in Figure 1-1 shows, it is likely that such an analysis would be performed by a specialist analyst because of the need to integrate a large number of often ill-fitting databases. These data sources can include vehicle registration details, previous victim databases, local crime information, prisoner release data (for release addresses of known offenders) and the location of local facilities such as psychiatric hospitals. Specialist software to perform this type of analytical function is available commercially, though at high cost. Although there is a need for the geographical profiling of offenders, the detection rate associated with these types of analysis accounts for less than 5% of recorded crime (Morgan and Newburn, 1997 p.118). The volume of these types of offences is low enough not to warrant a large number of analysts to be trained in this type of work, and therefore the model suggests that this type of work be performed by a specialised analyst.

High volume crime analysis

The analysis of high volume crime can be marshalled through a crime analyst and produces a number of desirable outputs:

Management information can aid senior and middle-ranking police managers when allocating officers to divisions and sub-divisional stations, and when deciding which areas will be the recipients of extra financial allocation for crime fighting.

Pattern analysis is the main function of a crime analyst involving the search for distinctive crime distributions. Hotspot analysis and identifying repeat victimisation come under this banner and the information gleaned from the system can be distributed within the force to aid policing.

Outside agency support is a growing area emphasised by the recent Crime and Disorder Bill 1998 which requires the British police services to co-operate more closely with outside agencies such as local councils and health authorities in order that a co-ordinated effort can be made to combat crime (Home Office, 1998). This support can include liaising with neighbourhood watch co-ordinators, crime prevention agencies, and improving public understanding of crime patterns through the newspapers or the use of internet web pages.

Street level support is shown as a support function for a crime analyst and is a possible direct output of a spatial crime analysis model. With the ability to customise modern GIS through programming languages such as Avenue and MapBasic it is possible to create a robust and user-friendly interface to a crime system which would allow police officers with no GIS experience to pursue their own lines of enquiry. Simple dialogs and managed output options allow officers to query the crime database and create their own maps of particular crime distributions. The support role of the crime analyst might be necessary to monitor usage and to give basic instruction. With skilful and sensible programming a robust 'officer-proof' system could be created.

Other areas of potential investigation shown are the analysis of traffic accident black spots and the addition of incident data into the model, though these would require additional data sets and analyst skills.

It is because of this wider array of potential applications within the high-volume crime area that this thesis will concentrate on the high volume analysis aspect of spatial crime analysis. It will converge on a number of areas of concern to the police: the accurate analysis of crime data, the identification and understanding of patterns of repeat victimisation, and the identification of crime hotspots. The importance of these areas for policing has been identified with the recent news that the government invited police forces, in consultation with local crime and disorder partnerships and police authorities, to participate in the targeted policing initiative of the Crime Reduction Programme (CRP) to start in April 1999 (Home Office, 1999). The initiative is part of the £250m crime reduction programme which the Home Secretary [Jack Straw] announced on 21 July 1998. The overall crime reduction programme will pilot and evaluate measures to build up evidence on what is effective and cost-effective in reducing crime. The project aims to combat the long term rise in crime by accumulating evidence about effectiveness including; targeting crime hot spots, repeat victims and repeat offenders. Access being unavailable to offender or criminal records, the chapters of this thesis discuss temporal analysis, and then concentrate on two of these three Home Office areas of concern: repeat victimisation and crime hotspots.

1.1.2. Perspective

The majority of crime analysis that takes place within the police happens at force headquarters where teams of analysts prepare management data for the highest echelons of the service and for the Home Office. This centralised analysis tends to focus on the preparation of statistical returns and in the interpretation of long term crime trends. The smallest spatial scale at this level is usually the police division and a force headquarters will rarely concern itself with sub-divisions or beats, leaving this level of analysis as the responsibility of the divisional commander. This thesis will examine the analytical possibilities from the viewpoint of a crime analyst at a divisional station where the change in scale requires a different suite of analytical tools and methods. At this divisional level useful spatial techniques are often unavailable, or in the case of some hotspot routines, inappropriate (hotspot algorithms are discussed in Chapter 8). The change in scale from the force-wide concern to the divisional and sub-divisional level also increases the need for spatial accuracy. A technique for minimising error in crime location is discussed in Chapter 7. This on-going theme of analytical techniques employable at a divisional and sub-divisional level runs throughout the thesis, and is particularly appropriate with regard to the recognised need for trained analysts in Nottinghamshire (HMIC, 1997a p.33). The national move from centralised analysis to the more proactive use of local intelligence resources within an integrated crime management model has been highlighted by two recent reports (Amey *et al.*, 1996; HMIC, 1997).

1.1.3. Nottinghamshire Constabulary

The primary source of data for this thesis has been the crime data recorded by Nottinghamshire Constabulary (this and other data sources are reviewed in Chapter 3). Nottinghamshire Constabulary records some of the highest levels of crime in the country, though some of this high recording rate has been attributed to the methods of crime reporting by the force (Nottinghamshire Constabulary, 1997 p.34). The force has remained consistently one of the highest crime areas in the UK for the last ten years. Figure 1-2 shows the levels of recorded crime in England and Wales for 1996. This general picture should be treated with caution as the police area level does not show important variations and differences within force areas (Maguire, 1997). Nottinghamshire recorded the highest per capita crime rate in the country in 1995 (Home Office, 1997), and this statistic is often held up as an example of the potential misleading nature of official statistics. A detailed study has shown that much of this high rate is due to different crime recording practices within Nottinghamshire Constabulary and that 'almost certainly, Nottinghamshire has never been the most criminal area in the country' (Farrington and Dowds, 1985 p70-1). Crime and incident data from Nottinghamshire Constabulary is discussed in greater detail in Chapter 3.

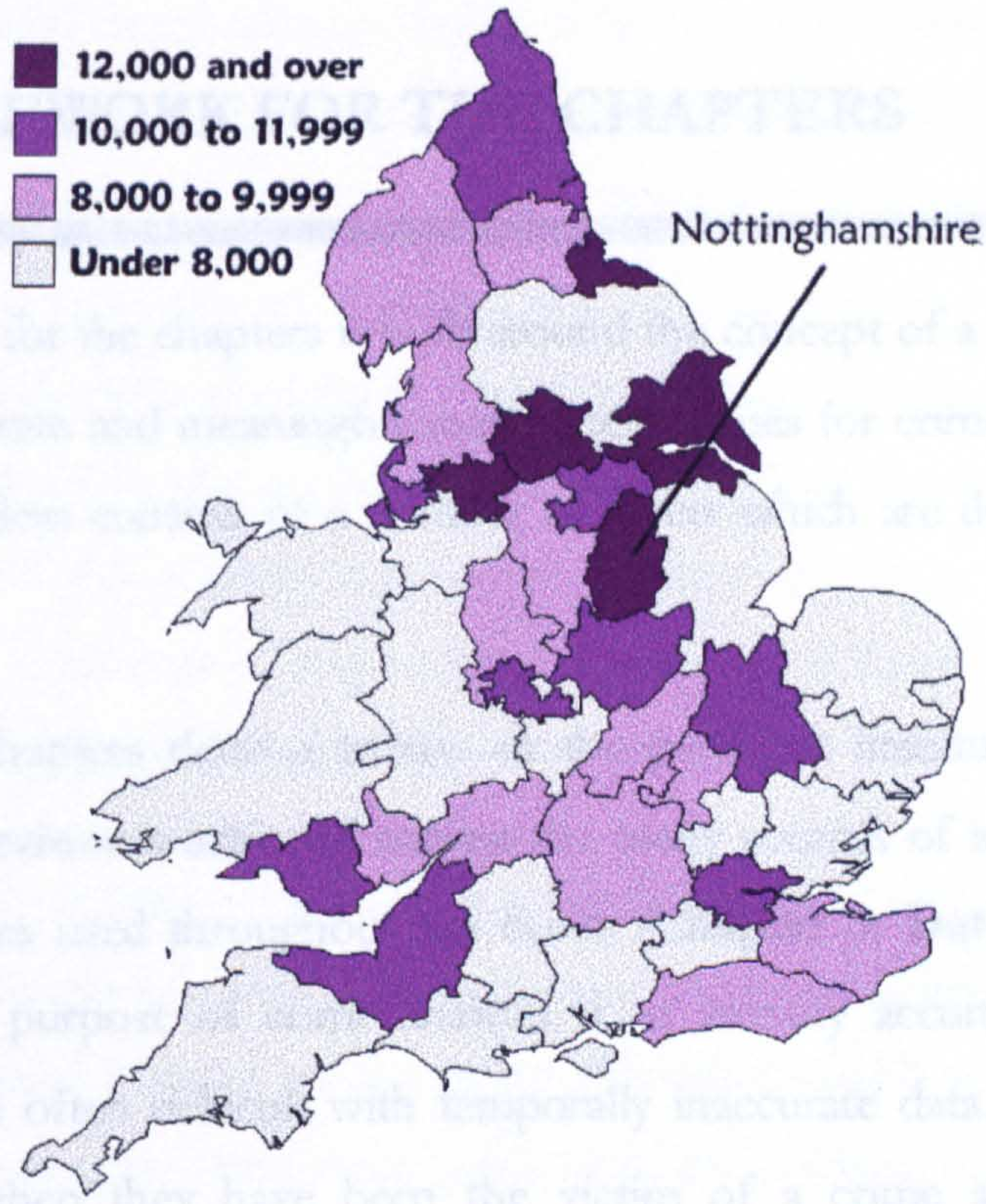


Figure 1-2 Notifiable offences per 100,000 population during 1996 for the police force areas of England and Wales.

Notifiable offences are generally (though not exactly) those offences which are considered more serious and triable in a Crown Court. Summary offences which can only be tried in a Magistrates Court are not included. Source: Criminal Statistics, 1996.

1.2. FRAMEWORK FOR THE CHAPTERS

The framework for the chapters is built around the concept of a single problem: the lack of accurate and meaningful analysis techniques for crime at a divisional level. This problem consists of a number of facets which are dealt with in the later chapters.

The first two chapters detail a review of the pertinent literature in this field (**Chapter 2: Previous work**) and outline the many sources of information and analysis processes used throughout the thesis (**Chapter 3: Data sources and software**). The purpose of crime analysis is to identify accurately significant patterns. This is often difficult with temporally inaccurate data. Victims rarely know exactly when they have been the victim of a crime and **Chapter 4 (Aoristic crime analysis)** addresses this issue. It introduces a new method to improve the accuracy of analysing temporally unspecific data. The application of this process in an example study area in Nottingham (UK) identifies a crime pattern not visible with other methods. The key aspects of this chapter have been published in the International Journal of Geographical Information Science (Ratcliffe and McCullagh, 1998). **Chapter 5 (Spatiotemporal differences between crime and incidents)** explores further the temporal dimension to crime with the use of seasonality analysis and correlation functions. This chapter also includes the use of incident data (recently made available by Nottinghamshire Constabulary) and explores the temporal and spatial relationship between recorded crime and incidents for assaults and disorder in Mansfield, Nottinghamshire.

The subject of **Chapter 6 (Identifying repeat victimisation)** changes the emphasis of the thesis to another area of policing concern. An investigation into the problems of identifying rapidly and accurately locations of repeat victimisation proposes that the use of GIS with georeferenced crime data can solve many of the difficulties identified. The findings of this chapter have been published in the British Journal of Criminology (Ratcliffe and McCullagh, 1998a). The spatial distribution of repeat victimisation is further investigated in **Chapter 7 (Burglary, victimisation and social deprivation)** by examining the spatial relationship with the socio-economic level in the vicinity of the offence locations. A new technique to aggregate polygon values in the vicinity of crime locations is introduced to alleviate the potential problem of georeferencing errors in the crime locations. An application using a deprivation index and burglary locations in South Nottingham reveals new insights into the spatial variation of burglary repeat victimisation, and has been published in Crime Prevention and Community Safety (Ratcliffe and McCullagh, 1999).

The final two analysis chapters investigate different processes for the identification of areas of crime concentrations, known as 'hotspots'. **Chapter 8 (Hotspot analysis)** reviews a number of currently available systems, some of which are crime-specific and others come from the fields of epidemiology and surface modelling. The chapter proposes a two-stage system which produces statistically supportable hotspots which are not restricted to certain geometric shapes but follow the morphology of the underlying crime events. Some aspects of this chapter have appeared as a chapter in a book by LaVigne and Wartell titled *Crime Mapping Case Studies: Successes in the Field* (Ratcliffe and McCullagh, 1998b). This process is applied in **Chapter 9 (Hotspots and police perception)** to a number of high-volume crime problems in South Nottingham. The hotspots are compared to a survey of police officers asked about their perception of high volume crime. The results have a number of implications for intelligence dissemination within the police service. Aspects of this chapter have been submitted to the Journal of Geographical Systems (Ratcliffe and McCullagh, submitted).

The final chapter (**Chapter 10: Conclusion**) briefly reviews the thesis in regard to the aims stated of developing techniques for divisional level analysis and draws a number of conclusions and recommendations for further work.

1.2.1. The accompanying CD-ROM and example data

The CD-ROM which accompanies this thesis contains a number of animations of time series data, examples of the software written to perform certain functions through the thesis, and some example data sets. It should be noted that the example data sets have been altered sufficiently to prevent the identification of any specific locations.

1.2.2. Summary

This chapter has introduced the reasons for completing this thesis and outlined the basic structure of the text to follow. The spatial crime analysis model identifies a number of areas where a divisional level crime analyst can contribute to the crime interpretation and prevention work taking place in police forces around the country, though the lack of analytical tools and methodology has been highlighted.

This thesis aims to investigate the GIS, spatial and temporal analysis techniques that can make a contribution to the analysis and prevention of crime. This is done from the perspective of a divisional crime analyst and concentrates on areas of current policing concern: accurate temporal analysis, repeat victimisation and hotspot analysis.

2. Previous work

This chapter introduces the main themes of the thesis and presents the significant literature surrounding those themes. The aim of the chapter is to critically review crime as a whole and identify the aspects of the subject where this thesis can make a contribution. Later chapters will review the most pertinent works and their contribution to the discussion and analysis that follows.

2.1. INTRODUCTION

The study of crime has traditionally been the preserve of other disciplines such as sociology, psychology and criminology (Georges, 1978). However since the 1970's there has been a realisation that there is a spatial aspect to crime, a study which is a legitimate subject by geographers. Crime occurrence can be modelled in a geographical context and a greater understanding of the patterns of distribution can be sought. Through a spatial analysis approach geographers hope to develop associational and predictive models that explain crime manifestation in regard to location. This type of work has a practical application in that the analysis of the dynamics of crime may assist those charged with the responsibility of crime control to assess better the effectiveness of programs they currently use, and to develop improved programs for the future.

> *The objectives of the geography of crime are to describe and map the spatial distribution of crime in greater detail and meaning than has been done before. This field of research attempts to relate the spatial patterns of crime to the environmental, social, historical, psychological (cognitive), and economic variables that may explain these patterns (Georges, 1978).*

Daniel Georges summarised the aim of geographers with respect to crime over the previous twenty years. While sporadic research successes have been recorded (mainly in the spatial relationship between crime and unemployment factors), it has subsequently been discovered that the causes of crime patterns are complex, and models of crime distribution have been few and far between (Herbert, 1976; Trickett *et al.*, 1995). The geography of crime covers a number of different topics within the discipline and the following chapter aims to review the main areas which have attracted research interest, and which are relevant to this thesis.

2.1.1. Crime and social factors

The first studies of the geographical incidence of crime and delinquency have been attributed to the French cartographer, Andre-Michel Guerry, whose pioneering work began with the 'Cartographic School' in the 1830's (Cater and Jones, 1992). Figure 2-1 is an example of the work of Guerry showing a map of crimes against property within the administrative regions of France first published in 1833 (Guerry, 1833). Work in France and in Britain immediately identified the relationship between crime and the urban-industrial region, and while there have been fluctuations in crime definition and data collection methodology, the spatial relationship of crime with poor inner cities has remained fairly constant.

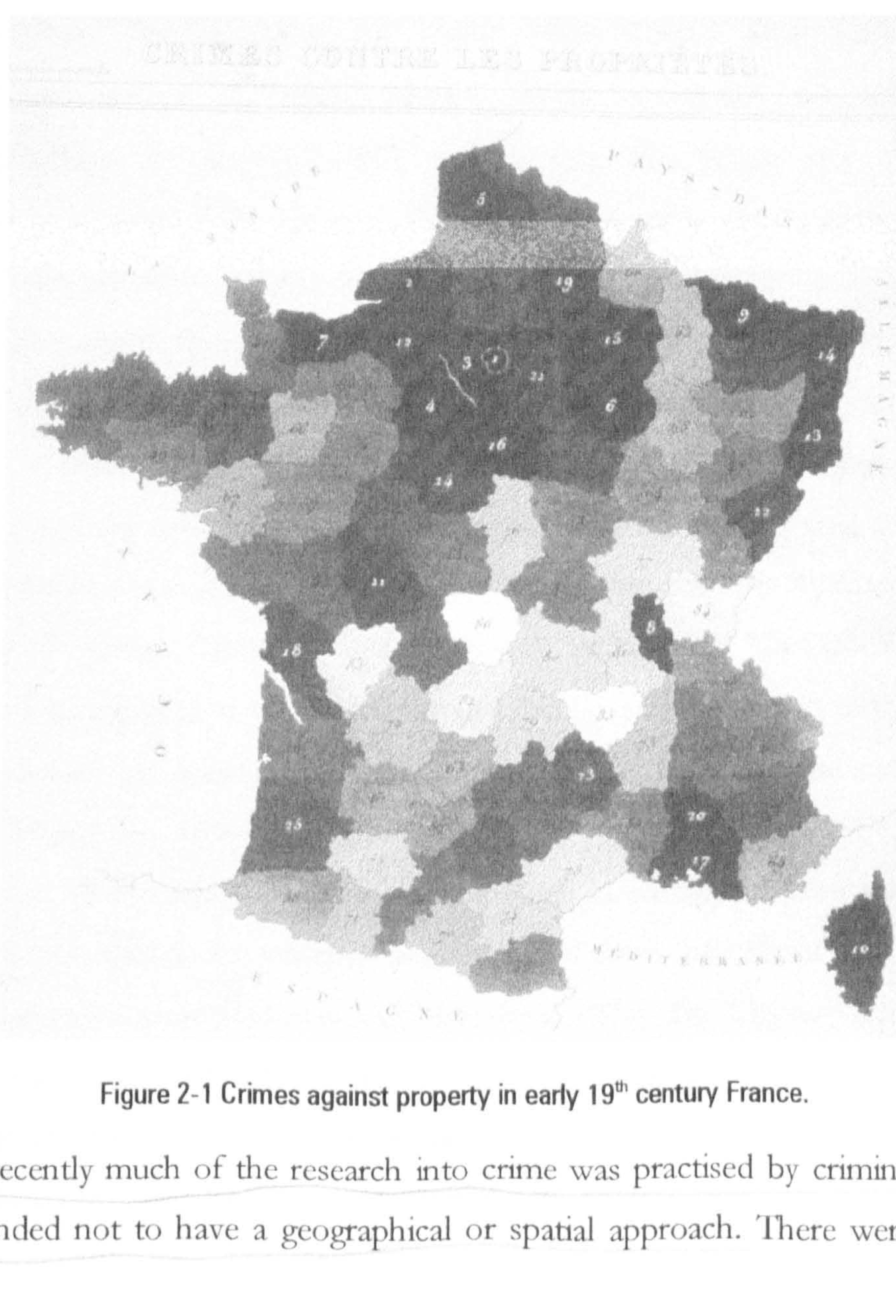


Figure 2-1 Crimes against property in early 19th century France.

Until recently much of the research into crime was practised by criminologists and tended not to have a geographical or spatial approach. There were some noticeable exceptions, especially in the relationship between crime and social factors. In one of the earliest UK spatial examinations of crime, Herbert attempted to examine geographically the cases of delinquency in Cardiff in the mid 1960's by focusing on the socio-economic causes and comparing these to the rate of juvenile offending (Herbert, 1976). Although there have been other examples, the lack of British research between crime and long-term unemployment in the UK has been recently highlighted with a call for more co-operation between the police and academic research (Amann, 1996). This desire for greater study within the UK may be due to the realisation that investigations into crime patterns have tended to be more advanced in North America than in Europe (Herbert, 1982). Examples abound and a number of references can act as a starting point (for example Allen, 1996; Anderson and Diaz, 1996; Brantingham and Brantingham, 1981; Cohen and Felson, 1979; Grogger and Weatherford, 1995; Kennedy and Veitch, 1997; Klinger, 1997; Ohlemacher, 1995; Sherman *et al.*, 1991). Hakim (1982) recognised the effects of unemployment on, amongst other things, crime and delinquency. Although written in a social policy journal, her article recognises the existence of the crime/unemployment relationship, with crime considered to be a knock-on effect of unemployment. Later work on the social consequences of crime sought to cement the relationship between socio-economic factors and crime. Brown (1982) performed a multivariate statistical analysis of crime in Chicago which found that there was significant clustering of crime in the ghettoised areas with little spill over into other suburbs - concluding that crime is localised in the ghettos of Chicago. Chicago was also home to the so-called 'Chicago School', a group of sociologists at the University of Chicago who sought to analyse their city on a socio-geographical basis by mapping factors such as crime and mental illness across the various zones of the city in the 1940's and 1950's. They combined official data with other research such as participant observation and life histories, though the patterns which resulted from the official data had the most research impact (Coleman and Moynihan, 1996). The Chicago School was the initiator of much of the spatial criminology which exists today. The ecological focus of their work can be related easily to the spatial analysis which is performed nowadays with the advantage of computers. One of the major works of the Chicago School was the mapping of crime occurrence showing a decrease radiating out from the Central Business District (CBD), correlating crime with Burgess' concentric-zone model (Shaw and McKay, 1942). This pattern was found to be spatially and temporally consistent with the highest occurrences of crime and criminal residence found adjacent to the CBD in; Chicago, Philadelphia, Cleveland and Seattle. The pattern in Chicago persisted through to the 1960's and beyond.

These studies confirmed the links between crime and delinquency and the social cohesiveness of which unemployment is a factor. A number of subsequent studies of a more overt geographical nature support this belief of a relationship between crime and unemployment (Reilly and Witt, 1992; Smith, 1986), though it should be noted that Reilly and Witt have been criticised for their mathematical methodology and statistical rigour (Pyle and Deadman, 1994).

The spatial pattern of crime is therefore not uniform and there are pockets of high crime concentration. In America these occur in the inner-cities, while in Britain; "(… with strong public-sector policies of slum clearance and rehousing) this simple geometry has been succeeded by a bi-polar pattern, with peripheral 'problem estates' of public-sector housing displaying relatively high levels of criminal activity." (Cater and Jones, 1992, p.81).

2.1.2. Crime and Space

Can a direct relationship between crime and space be found or is the relationship between space and other ecological variables such as class or poverty, with crime being a passive result of the segregation of those disposed to commit and be a victim of crime? One argument is that the link between poverty and crime is not deterministic but is complicated by the effects of a number of other variables such as demography (age, sex, family status), socio-economic status (income, education, unemployment), and living conditions (housing, density, tenure). Many of these variables are not necessarily linked directly to poverty. Space therefore becomes an active variable but not the sole cause of crime *per se*. Inner cities remain crime hotspots even though the socio-economic structure of the residents fluctuates over time. In addition to this, the Sheffield Crime Surveys of the mid 1970's found that lower class individuals were more likely to commit crime if they lived in a lower class community than the same type of people living in a mixed or higher class neighbourhood (Baldwin and Bottoms, 1976). Other relationships have included tenure (Baldwin and Bottoms, 1976), unemployment and social disorganisation. This suggests that mass urbanisation has broken down the social cohesion and traditional community bonds which instilled the population with a sense of community and moral values. Large urban settlements have induced a sense of anonymity that permits greater freedom of action and where criminality is less obvious. A recent example of this type of research has been completed in Liverpool (Hirschfield and Bowers, 1997) with a principal components analysis examination of the impact of social cohesion on crime rates. A general critique of this multivariate study is that the end product is entirely dependent on subjective decisions on the inclusion, exclusion and weighting of the chosen variables.

There is an alternative social theory to the general idea that delinquents and criminals are individuals who are outside mainstream society and are the result of inadequate socialisation. The Sub-Cultural Theory suggests that in fact these people have 'been socialised all too well into an alternative sub culture, many of whose values are diametrically opposed to mainstream morality' (Cater and Jones, 1992, p.88). Instead of criminals being a collection of individuals outcast from the main ideas of society, they might now be viewed as a separate group with its own cultural identity, i.e. the criminal fraternity. This possibility was noticed as far back as 1862 when Mayhew (1862) reported that children appeared to be 'born and bred to the business of crime'; this was echoed by Herbert (Herbert, 1982).

2.1.3. Travel to crime

Brown (1982) is an early example of a study into 'travel to crime'. Travel to crime studies tend to focus on the distance a criminal is prepared to travel to perpetrate a criminal offence and rely on arrest records of individuals or occasionally interviews with prisoners. Of local interest at this point is the PhD thesis completed by Jane Bradbury in the Geography Department of the University of Nottingham (Bradbury, 1981). Without the advantage of digitised data and therefore effective computerised analysis techniques for dealing with large volumes of data, Jane Bradbury had to amalgamate enumeration districts and wards in her study of greater Nottingham to enable her to cope with the mass of material. She examined the 'travel to crime' around Nottingham city in a study which identified that juveniles have less access to travel, committing crime within a relatively short distance from their home address. This finding correlates with the work of Brown, and is one of the few available academic examinations of crime in Nottinghamshire.

A number of writers have noticed a pronounced decay effect where the majority of property-related crimes occur near the criminal's home address, and decay rapidly the further you get from the home address (Bradbury, 1981; Brantingham and Brantingham, 1981). Another reason put forward for the decay effect in both adults and juveniles has been the suggestion of 'spatial awareness biases' (Cater and Jones, 1992). This suggests a behavioural pattern whereby factors of distance and information are relevant. A criminal will confine his activity to known areas (generally the local environment) and is 'unlikely to penetrate into totally foreign areas where he will feel uncomfortable or stand out as different.' (Brantingham and Brantingham, 1981, p.29). Evidence for this argument is compelling. A study in a Staffordshire town showed nearly 50% of detected burglaries were committed within 0.8kms of the perpetrators home (Cater and Jones, 1992). The Sheffield Crime Survey (Baldwin and Bottoms, 1976) found three quarters of the city's burglaries to have been committed within 2 miles of the burglar's address.

2.1.4. Police perception of crime

Though police crime reports are used as indicators of crime levels, criminal incidents reported to the police are not necessarily converted into crime reports. Police officers exercise their discretion at varying levels every day. If they are concentrating on one matter they may decide to overlook more minor transgressions in the hope of detecting a more serious offence. Many infringements of the road traffic acts such as speeding are often ignored if officers are *en route* to a more pressing matter. Furthermore, working class areas, and those with a reputation for crime, are more intensively policed, so that more crime is detected and the data may be biased (Nagle, 1995).

Two studies identified the affect that police officers had on the crime statistics (Bottomley and Coleman, 1976; McCabe and Sutcliffe, 1978). Bottomley and Coleman (1978) examined the role that the police play in the rates of recorded crime by examining crime reporting and the content of crime report forms. They also looked at clear-up rates and the problems of rate variation across police regions. They concluded that the emphasis of initiating a police response to crime should rest not on the police but on the public, as police play only a minor role (~13%) in the discovery of crimes which get entered onto crime reports. This figure corroborates a similar figure quoted by McCabe (1978).

In the second study (McCabe and Sutcliffe, 1978), researchers spent many months with officers from Thames Valley and recorded how the officers perceived crime. Most of the work discusses how police officers *define* crime by their decisions as to whether (or not) and how they record the crime. The officers have a significant effect on the crime figures depending on whether they record incidents as crime or not, arrest or not, and on what advice they give to members of the public. Such advice can range from reporting the incident as a crime, seeking a civil law solution, or simply advice on crime prevention. Often the advice can include 'this has nothing to do with the police'. It is interesting to note that while government may go to some length to define crime for the statute books, police officers redefine crime as it affects the area they police. This is also evident in a recent study from the United States (Klinger, 1997).

In an analysis of a number of American police precincts, Klinger's study explains that officers can exercise their discretion without interference from senior management (Klinger, 1997). The policing priorities of the patrol area contrive to ensure that officers in crime-ridden areas go to more important calls and therefore arrive at the notion that their area is a higher crime region than other beats. Officers arrive at their impression of the deviance of their area by observation and communicating informally with their colleagues. The result of this is that lower priority matters are dealt with informally and the sanctions of arrest and prosecution are reserved for the more serious crimes. Crimes which in many other areas would be reported and enter the criminal data system are therefore ignored. The relationship between crime area and police perception of crime location is examined in greater depth in chapter 9 (Hotspots and police perception).

2.1.5. Unusual studies of crime data

One of the more unusual uses of police recorded crime data has come from the group who analysed three years of monthly crime data on Merseyside to test the hypothesis that the Merseyside crime rate was reduced by a group practising Maharishi Mahesh Yogi's 'Transcendental Meditation' (Hatchard *et al.*, 1996). Apparently a 3.4% drop in crime was recorded when the group size first exceeded the 'root 1%' or Maharishi Effect threshold! Possibly spurious relationships such as this (another example might be the correlation between the increase in crime in the 1980s and the growth in the use of mobile telephones) hide more unusual papers which have found surprising relationships with crime. An analysis of annual, quarterly, and monthly data for recorded crime in England and Wales, yielded strong evidence that temperature has a positive effect on most types of property and violent crime (Field, 1992). The author hypothesised that this effect was due to people spending more time outside in the good weather, both leaving their property unguarded (increasing property crime) and participating in more social interaction (increasing violent crime).

2.2. CRIME DATA

2.2.1. Limitations of crime data

The main problem highlighted within many studies is the inherent inaccuracy of crime figures. It is extremely difficult to get an accurate picture of the level of true crime in an area. Early and current discussions surround the merits of the various data sources, which include; court records, crimes recorded by the police, household and crime surveys, and offender interviews. This is not a new problem.

Thirty-two years ago Biderman and Reiss reviewed the discussions surrounding the imperfections in official data (Biderman and Reiss, 1967). Data on crime has to be sourced somehow, and the paper identified the problems with police data, court records and other official sources of information on crime. While it was known that this 'institutional' data did not reflect a true picture of the volume of crime and ignored unreported/undetected incidents (Biderman and Reiss' 'dark figure' of crime), the authors discussed two of the differing traditional responses to this; the 'realist view' and the 'institutionalist view'. Realists wanted to see the 'actual' amount of crime - and so uncover the 'dark figure' of hidden crime, whereas institutionalists emphasised that crime could only have a valid meaning in terms of the organised social response to crime. The institutionalist view was that a coherent response to crime was only possible in the context of the reported/recorded crime. They urged the study of the official criminal records as outcomes of social and institutional processes (Coleman and Moynihan, 1996). However, the realists' view has been adopted this century with police records forming the basis of crime statistics since 1923 (Biderman and Reiss, 1967). Having said this, realists do not doubt that figures produced by official crime statistics give only 'a partial view of crime in society' (Smith, 1986). Biderman and Reiss approached the subject with the aim of using official statistics to analyse crime in a social context. Although not strictly geographical, but more of a historical account of the discussions around crime statistics, it is useful to note that problems of crime data collection identified more than thirty years ago still have not been successfully resolved.

There have been a number of attempts to correct for the under-reporting of crime. Susan Smith details the recording of various crimes from non-police sources instead of police crime records, finding these provide a fuller account than the police records. This research included: telephone box vandalism from Post Office engineer reports, assaults on bus drivers from transport authority reports, and malicious false alarms from fire service operator sheets (Smith, 1986). Unfortunately this type of data recording method tends to be useful only for specific localised applications. On a broader scale the British Crime Surveys have been the best attempt in recent years to gather an accurate picture of the 'dark figure' crime levels across the country.

2.2.2. The British Crime Surveys

The realisation that not every crime becomes a statistic has lead to various attempted solutions of which the British Crime Survey (BCS) is one. The BCS uses extensive survey information to attempt to elicit a more realistic estimation of national crime and discover more about the 'dark figure' (Hough and Lewis, 1989). The biannual survey asks a large representative sample of the general public about their experience as victims of household and personal crime in the previous year. Its main aim is to provide a count of crime that includes incidents not reported to the police and is therefore not affected by changes in the way police record crime. This means that the sampling technique is fairly uniform across the survey and removes the effect that the police have on the crime statistics as mentioned earlier (Bottomley and Coleman, 1976; McCabe and Sutcliffe, 1978). It therefore provides a more complete picture of the national extent of crime than police figures, for the offences it covers. It also gives a better measure of the trends in crime, as willingness to report crimes to the police varies depending on the type of crime (Table 2-1).

Table 2-1 Reasons for not reporting crime: 1996 BCS

Reason for not reporting crime to the police	%
Too trivial / no loss involved	40
Police could do nothing	29
Police would not be interested	20
We dealt with matter ourselves / inappropriate for police	19
Reported to other authorities	5
Inconvenient to report	4
Fear reprisals	4
Fear / dislike police	<1
Other	5

Source: Mirrless-Black et al., *1996, p.63.*

The British Crime Survey (Mirrless-Black *et al.*, 1996) estimated the following statistics. We can expect: a burglary in the home every 40 years, a family car to be stolen every 60 years, an assault resulting in injury once every century, and a robbery once every 500 years. Besides counting crimes, the survey can also show which types of people are most at risk of which types of crime and describe the nature of crime: where, when and how it happened. It is also used to ask people about other crime-related issues, such as fear of crime and attitudes towards the police. The Home Office has now conducted six sweeps of the survey (1982, 1984, 1988, 1992, 1994 and 1996). In 1996 more than 16,500 people aged 16 and over were interviewed.

The British Crime Surveys reinforce the point that analyses based on police statistics do not necessarily reflect the real level of crime: only about a quarter of all crimes are reported to the police (Mayhew *et al.*, 1993). Many crimes of an ingenious nature such as tax evasion, computer crime and subtle cases of arson are not detected by the public or police and escape reporting. Many offences are not reported for personal reasons. Sexual offences may cause embarrassment to the victim, and to the victim's family in cases of incest. When the victim is wanted by the police they often do not contact the police for fear of being taken into custody, and shoplifters may just have the stolen goods confiscated with out the police being contacted. Other reasons for non-reporting include: apathy, the notion that the police are either too busy or not interested, or the belief that the crime is not serious enough to warrant the involvement of the authorities. These attitudes are summarised in Table 2-1.

A point to note from the 1996 BCS is that the estimated reporting rate of burglary and motor vehicle theft to the police is over 80% (Table 2-2), whereas other recordable offences have much lower reporting rates (Mirrlees-Black *et al.*, 1996). A similarly high rate of over 90% was recorded in 1992 (Mayhew *et al.*, 1993). Many of the UK police forces and divisional priorities include reducing the levels of burglary and motor vehicle theft within their areas.

The most recently published survey from 1996 found that under half of BCS offences were reported to the police. Reporting rates varied by offence. Car thefts, burglaries and bicycle thefts were usually reported, while vandalism, thefts of property and thefts from the home often were not (Figure 2-2). It is likely that some of the variation in Figure 2-2 can be attributed to the benefits (or not) of reporting some types of crime. Insurance claims require a police crime reference

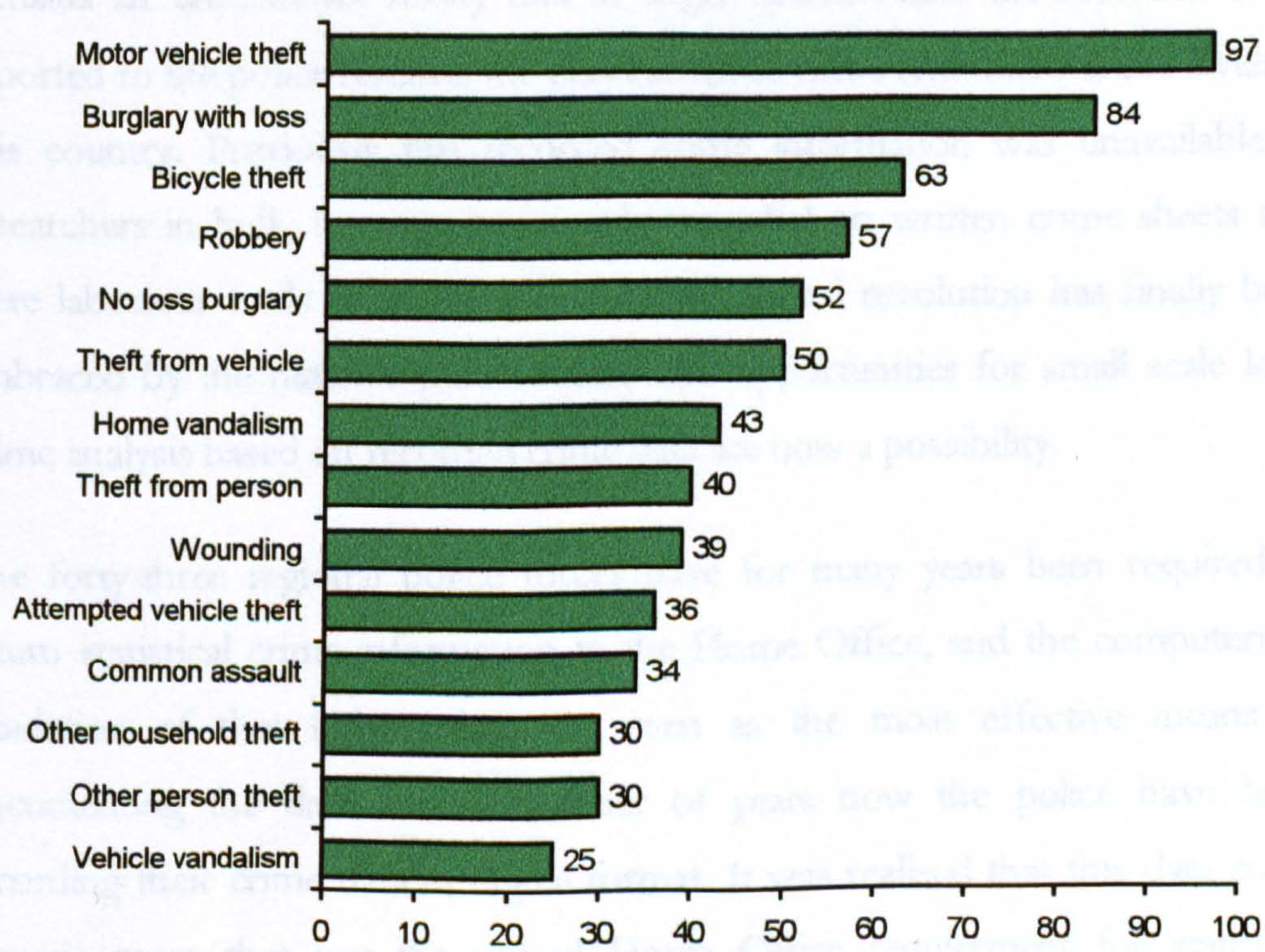


Figure 2-2 Percentage of incidents reported to the police (1996 BCS).

Source: Mirrless-Black et al., *1996, p.23.*

number and this could explain the high reporting rate for vehicle theft and burglary with loss. The 1996 BCS recorded falls compared to previous years in a number of crime categories for the first time. Since 1993, burglary with loss has decreased 5% in the BCS and 12 % in recorded crime figures. Reductions in car crime were also detected between 1993 and 1995 for the first time. Theft of vehicles and thefts from vehicles fell by 8% and 2% respectively in the BCS, with larger drops recorded in the police figures (Mirrlees-Black *et al.*, 1996).

While the British Crime Survey claims to have improved the prediction of 'actual' levels of crime, this has been at the expense of spatial resolution. The very nature of the survey population size means that the survey is only truly effective at the macro scale. This point was borne out by one publication which found significant regional variation in crime pattern trends by comparing the 1982 and 1988 British Crime Surveys (Trickett *et al.*, 1995). Changes in regional crime rates fluctuated in a complex manner at the regional level indicating that a number of elaborate forces were at work affecting the regional rates. This suggests the possibility that at even smaller study scales crime rate variation might be even more noticeable. This can not be measured by the BCS. It still remains an unfortunate reality that at larger spatial scales the recorded crime reported to the police remains the best record of meso and micro crime levels in this country. Previously this recorded crime information was unavailable to researchers in bulk, because it was only recorded on written crime sheets that were laborious tools to utilise. However the digital revolution has finally been embraced by the nation's police forces and opportunities for small scale local crime analysis based on recorded crime data are now a possibility.

The forty-three regional police forces have for many years been required to return statistical crime information to the Home Office, and the computerised tabulation of that information was seen as the most effective means of documenting the data. For a number of years now the police have been recording their crime data in digital format. It was realised that this data could provide more that just the annual Home Office requirement for statistical information. At the same time, Geographical Information Systems (GIS) began appearing in police stations, mainly tasked with complementing the immediate response nature of policing (responding to 999 calls) in command and control systems.

Geographical Information Systems have been quickly adapted by police forces for their use in a variety of different operational situations, though their use has not yet been fully utilised for crime mapping or analysis. There has recently been a flood of publicity as police forces both here and in the United States have invested in GIS technology (see Berkeley, 1997; Campbell, 1992; Clegg and Robson, 1995; Fox-Clinch, 1997; Grescoe, 1996; Hirschfield *et al*, 1995a; ICL, 1995; MapInfo, 1997; Mitchell, 1997; Nagle, 1995; Page, 1997; Salinas, 1997; Tempe, 1997). These systems have been used mainly for mapping live incident data, occasional crime mapping, and for describing incident scenes to a court. For example, in a Computer Aided Despatch (CAD) system, also known as a Command and Control system, the locations of live incidents from the emergency telephone network (999) can be displayed on screen enabling the operator to allocate efficiently the resource manpower most effective to deal with the particular event. It also assists the operator in visualising the scene and permits them to control better the resources 'on the ground'. Accurate guidance of responding officers to the scene of the incident can save valuable time. These same principles have also been applied to other emergency services (Smith, 1997). Most of the crime-related articles (in magazines and journals such as Mapping Awareness, Police Review and Geographical Magazine) discuss the *future* application of GIS for crime pattern analysis as a *future* development which the companies and organisations hope to develop.

2.3. CRIME ANALYSIS

Campbell (1992) identifies a number of areas in which GIS can contribute to the efficiency of police services including; information for evaluating and implementing changes in police patrol areas, manpower allocation at large public events, proactive policing strategies and contingency planning for major disasters. Chief Inspector Campbell also identified three key GIS objectives for Northumbria Police which related to crime mapping, and in doing so suggests a suitable 'wish-list' for police forces generally:

- Provide all officers with easy access to information which will assist them in the better performance of their duties;
- Provide a management tool to aid with better decision-making and the best possible deployment of resources;
- Provide a cost-efficient means of integrating databases.

During the early 1990s the first research papers began to appear which utilised the geocoding of police reported crime information (examples include Berry and Jones, 1995; Campbell, 1992; Ekblom, 1988; Grogger and Weatherford, 1995; Wrighton, 1987). The majority of examples were drawn from North America (Berkeley, 1997; Grogger and Weatherford, 1995; Maltz *et al.*, 1991; Salinas, 1997; Tempe, 1997), where the usage of GIS for crime mapping has been quite significant. Simple analysis and mapping have been the main applications (Figure 2-3). Ekblom, writing on the subject of crime analysis for the Home Office, considered; 'Operations include correlation, cluster analysis, regression and loglinear modelling. In most crime analysis exercises these are unlikely to reward the extra effort involved.' (1988, p.32). It should be noted that at the time of publication, spatial analytical toolkits were few and far between and computer processor speeds were slow enough that even minor iterative computations were time consuming. Berry and Jones' article (1995) carries an excellent discussion of police and crime census statistics. They used the postcode address file (PAF) and ARC/INFO to develop a crime based GIS, and concluded:

> *We anticipate that the single greatest potential for a comprehensive crime-based GIS is in the monitoring of any crime displacement, geographically, that is consequent upon spatially targeted crime-prevention measures. (p.77)*

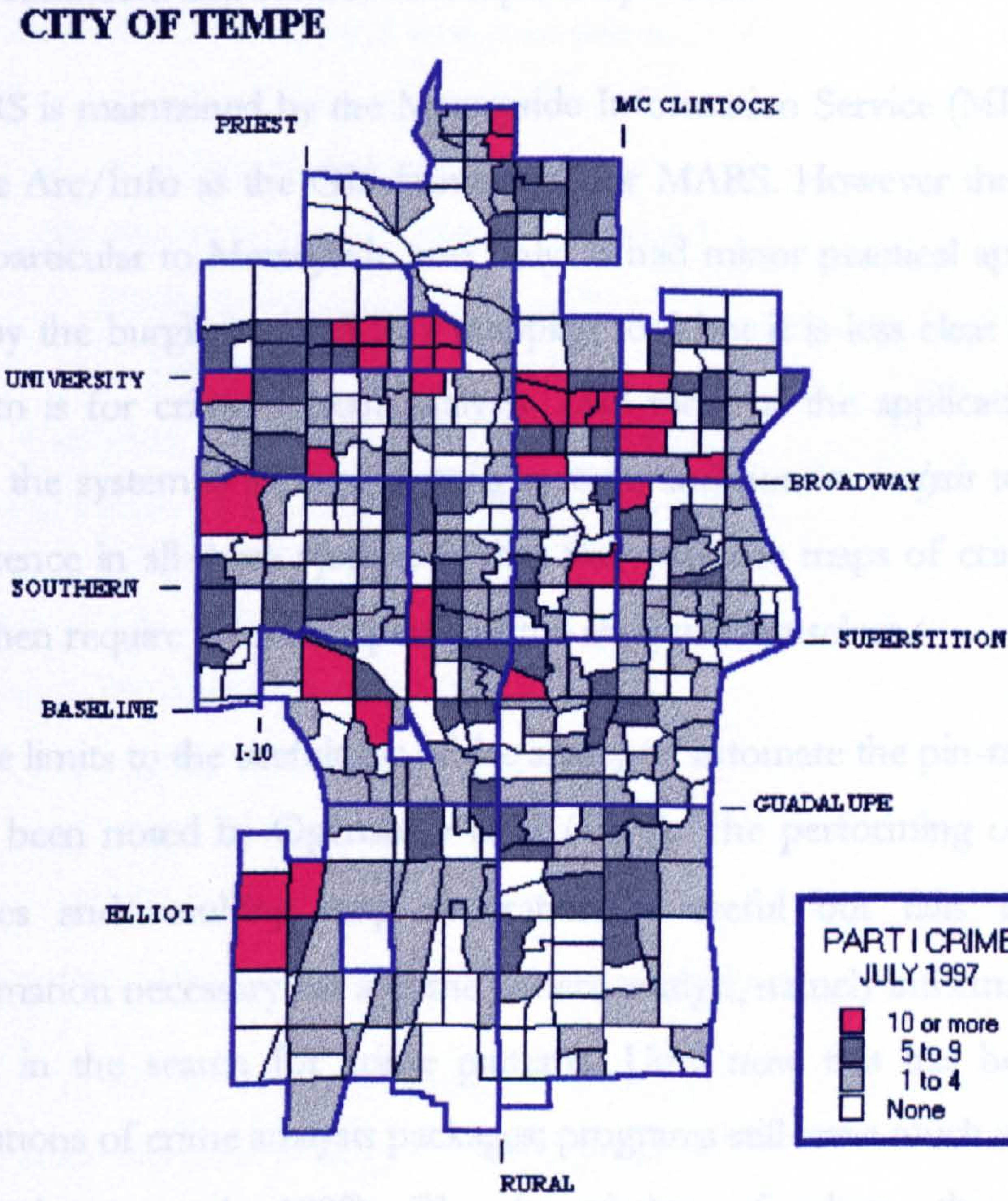


Figure 2-3 Example online map from Tempe Police Crime Analysis Unit home page.

Source: http://www.tempe.gov/police/

One group of researchers at Liverpool University (Hirschfield *et al.*, 1995a) have been developing a GIS-based crime analysis and mapping system for use in the analysis of crime incident data recorded by Merseyside Police. This has been developed using resources from the British Urban Crime Fund. Their decision to use GIS was prompted by a number of factors unique at the time to Merseyside. The region has a comprehensive road network and property database called the

Merseyside Address Referencing System (MARS), which forms the basis of the Command and Control and Crime Incident reporting system used by Merseyside Police. This database was available to the researchers, as was that part of MARS which provided a one-metre resolution grid reference for each property in the study area. Each property was identified by a Unique Property Reference Number (UPRN). The UPRN is recorded on every crime report and during their command and control data capture operation.

MARS is maintained by the Merseyside Information Service (MIS) who in 1991 chose Arc/Info as the GIS framework for MARS. However the MARS system was particular to Merseyside area only. It had minor practical applications in its use by the burglary squad as a mapping tool but it is less clear how useful the system is for crime pattern analysis. Like most of the applications mentioned here, the system is a crime *mapping* system, and not an *analysis* tool. The crucial difference in all these systems is that they produce maps of crime distribution, but then require the *user* to perform the analysis themselves.

These limits to the usefulness of the ability to automate the pin-mapping process have been noted by Openshaw *et al.* (1990). The performing of simple spatial queries and resulting map generation is useful but fails to provide the information necessary for a crime pattern analyst, namely information which will assist in the search for crime patterns. Until now this has been one of the limitations of crime analysis packages; programs still leave much work to the user (Openshaw *et al.*, 1990). The knowledge of where the highest crime concentrations are located is important to policing and crime prevention strategies, yet once a mapping system produces a number of maps, the human operator must still interpret the images and look for clusters or 'hotspots' of crime. One of the landmark papers in the area of crime mapping with GIS is the account of a *failed* early attempt to introduce a GIS system to Northumbria Constabulary in the late 1980's (Openshaw *et al.*, 1990). The paper raises a number of relevant issues by recounting the attempt to implement a GIS package into a sub-divisional police station in Northumbria. The project eventually failed and the reasons for this were documented by the author (Table 2-2).

Table 2-2 Some of the reasons identified for the failure of the Northumbria Police GIS.

Description of problem	Probable cause
Mapping of data was incomplete	software
Crime records could not be deleted	software
Data input too time consuming	customisation, design
Nothing could be retrieved once input	software, customisation
Method of handling information too complex	customisation
System not user friendly and accepts erroneous entries	customisation
No on-site support available	vendor
No comprehensive documentation	vendor
No training in its use, operators self-trained	vendor support
Plotter hardware interface problems	vendor support
User's lack of faith in vendor	vendor
User's loss of moral once vendor lost interest	user
Use of the system achieved nothing, seen as a burden, failed to live up to expectations	user
Plotting too slow	hardware
Database not suited to end user use	system design
System specification inadequate	inexperience
System failed to match expectations	vendor
No funding to make changes to system	vendor and user
System not fault tolerant but punished user for making errors	system design
OS map base 10 years out of date in places	data supplier
System supposed to be up and running on day 1	vendor
System could not be relied on	vendor

Source: (Openshaw et al., 1990).

A number of the problems identified are still relevant today even with the improvements in GIS commercial packages now available. Finding an academic solution to a problem lacks value if the means to implement that solution are not available or have not been thought out. Similar examples occur elsewhere.

The book by Maltz and his colleagues describes in great detail some of the problems of implementing a GIS crime mapping system for the Chicago police force (Maltz *et al.*, 1991). Obvious comparisons exist with the Northumbria article (Openshaw *et al.*, 1990) as in both cases the most difficult hurdles to overcome were often the problems of bringing change and a new technological approach to 'institutionalised' systems (the respective police forces); systems which had an enormous capacity to resist change. Lack of technical experience and knowledge of the end users was matched by professional cynicism on the part of the more established police officers.

2.3.1. Crime analysis systems

Any discussion of crime analysis systems has to begin with a definition of crime analysis, local crime analysis, and crime pattern analysis. There has been little formal attempt to define these terms, especially from the commercial companies who often use the terms with no explanation of the difference, even when they are aware of such a difference. The terms 'crime pattern analysis' and 'crime analysis' are often employed as if they have clearly understood meanings; indeed they are often used interchangeably. The TREVI sub-committee, designed to provide a European standard, identified eight different forms of crime analysis (Read and Oldfield, 1995):

- Crime Pattern Analysis
- General Profile Analysis
- Crime Control Methods Analysis
- Case Analysis
- Comparative Case Analysis
- Offender Group Analysis
- Specific Profile Analysis
- Investigations Analysis

There are two that are of specific relevance to this text. *Comparative Case Analysis* seeks to detect similarities between crimes that point to the same perpetrator or groups of perpetrators, and *Crime Pattern Analysis* which attempts to form a picture of the nature and scale of crime in a particular area. When these types of analysis are applied to the investigation of high volume crime at a local (either divisional or sub-divisional) level, then they come under a general term - *Local Crime Analysis* (Read and Oldfield, 1995).

COMPARATIVE CASE ANALYSIS

Crime Pattern Analysis (CPA) is one of the more misunderstood terms in the policing and criminological world. Many systems claim to perform *crime pattern analysis*. However, when looked at more closely, they are often found to be different beasts altogether. A number of different text-based search engines are employed by some police forces for major incident investigation. They work by searching large databases looking for links between one crime and another. If you enter 'blood-splattered kitchen knife' they will go to the database and return with every reference to 'blood', 'splattered', 'kitchen', and 'knife' (more or less). Although these systems deal with crime, all the system is doing is searching for matching similar strings of text. There is a distinctive lack of analysis taking place and the onus is still on the user to do all the thinking and come up with the right choice of search string. Without the capability to analyse structurally the material, they remain merely clever database managers. The output is geared towards understanding the result of the text search. These types of system are better described as Comparative Case Analysis systems. Commercial examples include the Harlequin Criminal Intelligence System (Harlequin, 1997) and ProQuest by i2 software (i2, 1997).

CRIME PATTERN ANALYSIS

Crime Pattern Analysis (CPA) can only be performed by systems which map crime distributions and then analyse these maps for significant patterns. At the simplest level this is the ability to examine a number of different areas, such as police beats for example, and identify which has the highest level of a certain type of crime. If this is performed by a mapping or GIS package, which can graphically display the output on a map, then this will improve the comprehension and quality of the analysis. If one has patience (and a sense of humour) one does not even need a computer for true CPA (though they do help when you are examining thousands of burglaries!).

A hotspot is an area of higher crime concentration where there is a greater density of criminal activity. At a more complex level an ideal system would be able to detect crime hotspots from the data by examining the whole area and focussing on the highest concentrations of crime. It should be able to do this irrespective of where boundaries such as police beats are located. Criminals have an annoying tendency to ignore police station demarcation areas, and on occasion they make deliberate use of county boundaries by crossing from one side to another to commit their crimes in an attempt to confuse police detection techniques (Williamson, 1998). Just like a GIS, the defining trait of a CPA system is the ability to perform specific spatial analysis techniques, ignoring boundaries when required. Unfortunately as we will see later, there are few systems which can do this at all, and those that do have significant limitations. The importance of crime hotspots can not be underestimated. With the increases in crime over the last 20 years, and the reduction in the numbers of operational police officers, it is vital that officers are directed to the worst areas.

SPATIAL ANALYSIS TECHNIQUES FOR HOTSPOT DETECTION

Crime events can be thought of as a continuity of instantiations over a study area. Much of the work has gone into searching this mass of points for distinctive clusters of crimes. The heaviest clusters signify areas with spatial clusters of criminal activity. This search for the heaviest crime clusters has been termed *hotspot analysis* (Hirschfield *et al.*, 1997).

Most of the literature relating to spatial crime analysis techniques share similarities and origins with aspects of medical geography and the search for clusters of rare diseases within medical epidemiology, research that goes back to the classic research of Knox and Mantel in the 1960's (Knox, 1964; Mantel, 1967; Openshaw *et al.*, 1988). One of the seminal works in the literature has been provided by Openshaw and Charlton (Openshaw and Charlton, 1987), in the form of the 'Mark 1 Geographical Analysis Machine' which sought to provide a fully automated process for the complex analysis of point pattern data. The technique involved four main steps;

1. Define a two-dimensional grid lattice over the study area and initial test circle sizes.
2. Over each grid intersection, apply a variety of different circle sizes and retrieve the events within the circle. Perform some statistical test on this data and store the results which achieve the given statistical level.
3. Increase the test circle size and change the grid mesh to reflect the new circle size.
4. Repeat steps two and three until the complete mesh has been tested with all relevant test radii.

Although attention is focused on a circular search region, the traditional geometry used in epidemiological studies, search regions such as hexagons have been used in other studies (Hirschfield *et al.*, 1997). Openshaw (1987) used a Monte Carlo simulation based on a simple count of the number of events within the circle. This test compared the observed count with the number that would be expected under a specific null hypothesis. Other papers have adapted the basic technique to different analysis methods such as quartic kernel intensity tests (Gatrell *et al.*, 1996). Again this application was for epidemiological studies.

Crime-specific systems have been few and far between. The Illinois Department of Justice have developed a system called the 'Spatial and Temporal Analysis of Crime' software (STAC) which is used for crime profiling in Chicago and elsewhere in the United States (Hirschfield *et al.*, 1997). STAC uses similar techniques to the Mark 1 GAM in that a grid of circles is searched for crime patterns. Spatial clusters of crime, termed 'Hot Clusters' are then identified. STAC summarises the boundaries of the defined clusters as standard deviational ellipses and the details are exported as a grid file of co-ordinates. The mathematical properties of standard deviational ellipses are detailed by David Ebdon (Ebdon, 1996).

One of the main research groups in the field has been active in converting the STAC system for use within a GIS. Alex Hirschfield and colleagues, based in Liverpool, have been working at Liverpool University in the Department of Civil Design on developing a GIS-based crime pattern analysis system as part of an ESRC-funded research project. Recently the group have been comparing different methods of spatial pattern analysis in an attempt to define the most suitable system for analysing their previous data. Currently no system has emerged as an outright winner (Hirschfield *et al*, 1997). Although they have converted STAC to run within a commercially available GIS, MapInfo, they have recorded a number of limitations and problems with this approach to crime pattern analysis. The system is unable to establish the statistical significance of the ellipses by comparing an observed number of incidents with an expected frequency, identify any variations in levels of crime within the hotspot, and there is a lack of correspondence between the shape of the hotspot and the underlying patterns of landuse (Hirschfield *et al*, 1997).

The STAC system is also susceptible to the Modifiable Areal Unit Problem (MAUP) (Hirschfield *et al*, 1997). The spacing of the grid and the size of hexagons which the system uses can generate completely different descriptions of the same crime data. In an attempt to investigate the problem the researchers moved onto developing the Mark 1 GAM for their system. To address the MAUP the analysis was run a number of times for many different sized circles. It was argued that real clusters would be manifest at all spatial scales. With limitations, their analysis would seem to bear this out. Chapter 8 (Hotspot analysis) describes the Modifiable Areal Unit Problem in more depth.

STAC is popular in the United States but is relatively unknown in the UK where there are few genuine crime analysis systems. Vertical Mapper is an example of an add-on program designed to integrate with MapInfo. It provides a mapping technique for data which vary continuously over a geographical area (Northwood Geoscience, 1998), and is mainly used for limited surface modelling. It has been used by the Brent Crime Mapping Project in their mapping of the distribution of various crime types in a study area at Wembley, North London (BCMP, 1997). The output is similar to that produced by most surface modelling packages, and an example is presented in Figure 2-4.

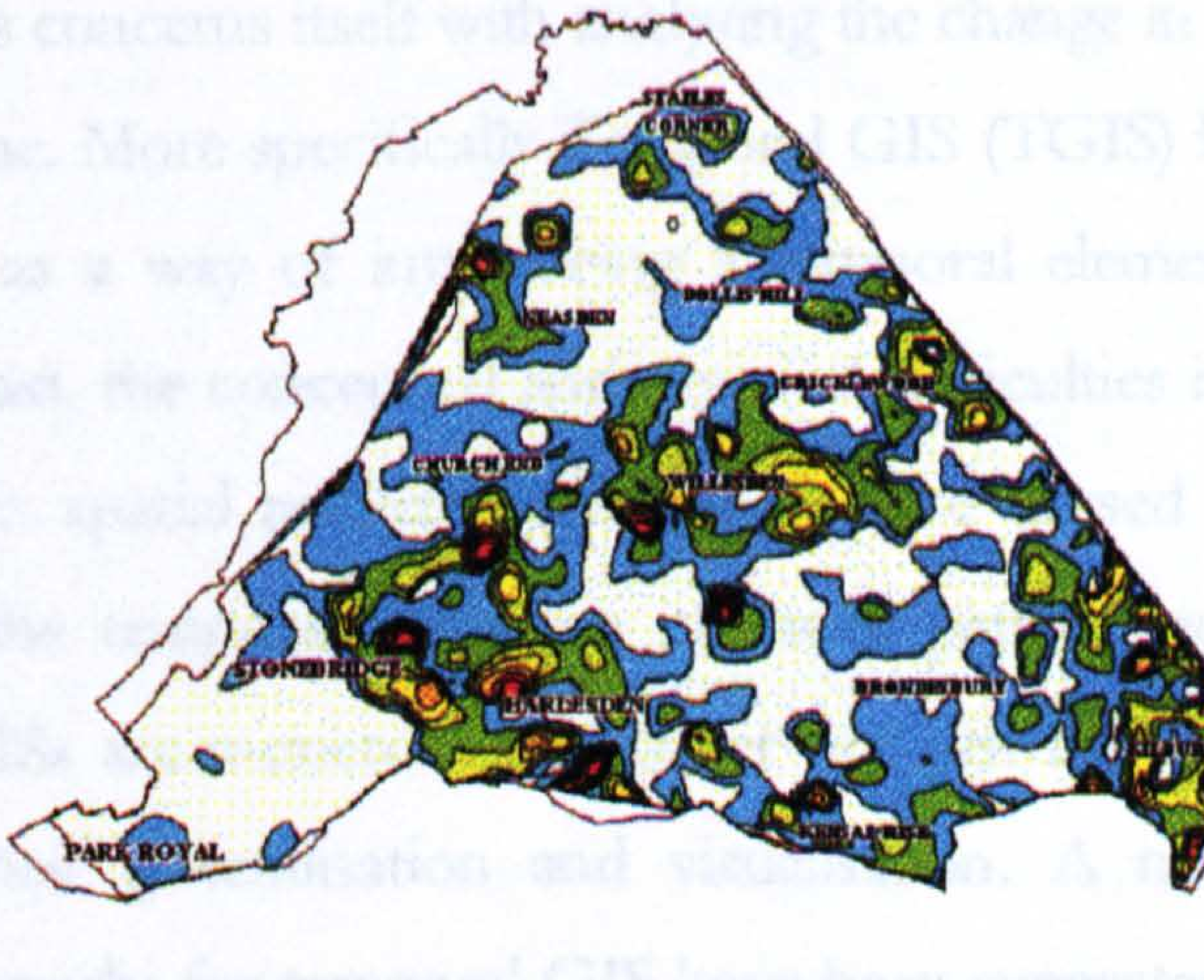


Figure 2-4 High volume crime in Harlesden June to November 1996, displayed using Vertical Mapper.

Source: Crime Awareness News, February 1998, Issue 6 (BCMP, 1998). This image shows a map of crime hotspots generated by Vertical Mapper. White and pale areas show regions with a negligible crime rate. As the colours progress through blue to green, yellow and red the level of crime incidence increases.

The Liverpool group have been focusing on the comparison of crime levels with the underlying population topography. There are limitations in this approach in relation to practical policing. While this type of analysis is of clear academic interest, a number of discussions which the author has held with police officers has identified the need for a simpler system which identifies areas with a crime problem. Incorporating the underlying population structure may shift the focus onto sparsely populated areas with moderate crime levels, and away from dense urban areas with higher levels of crime.

STAC and Vertical Mapper will be evaluated in greater depth in Chapter 8 (Hotspot analysis) where an improved technique for hotspot detection will be presented.

2.3.2. Temporal analysis

One possible application of crime data which has not been discussed in depth in the literature is the use of temporal data to plot changes over time. While many police officers are reasonably aware of where the maximum crime areas are on their beat, they may be interested to know when the picture of crime distribution changes. This can be done by plotting change over time.

Temporal analysis concerns itself with analysing the change in a variable or set of variables over time. More specifically Temporal GIS (TGIS) has been proposed (Peuquet, 1994) as a way of introducing a temporal element to geographical analysis. In the past, the conceptual and practical difficulties in representing and analysing complex spatial patterns within GIS have caused the representation and analysis of the temporal dynamics of those patterns to be ignored. The problems of TGISs are numerous and cover not just the analysis of temporal data, but also their generalisation and visualisation. A number of different conceptual frameworks for temporal GIS have been suggested (see Peuquet and Niu, 1995; Raafat *et al.*, 1994) but a defining standard has yet to be agreed. The full potential of space-time models for empirical spatial analysis thus remains unrealised because the paradigm for representing data temporally in the manner required does not yet exist (Miller, 1991).

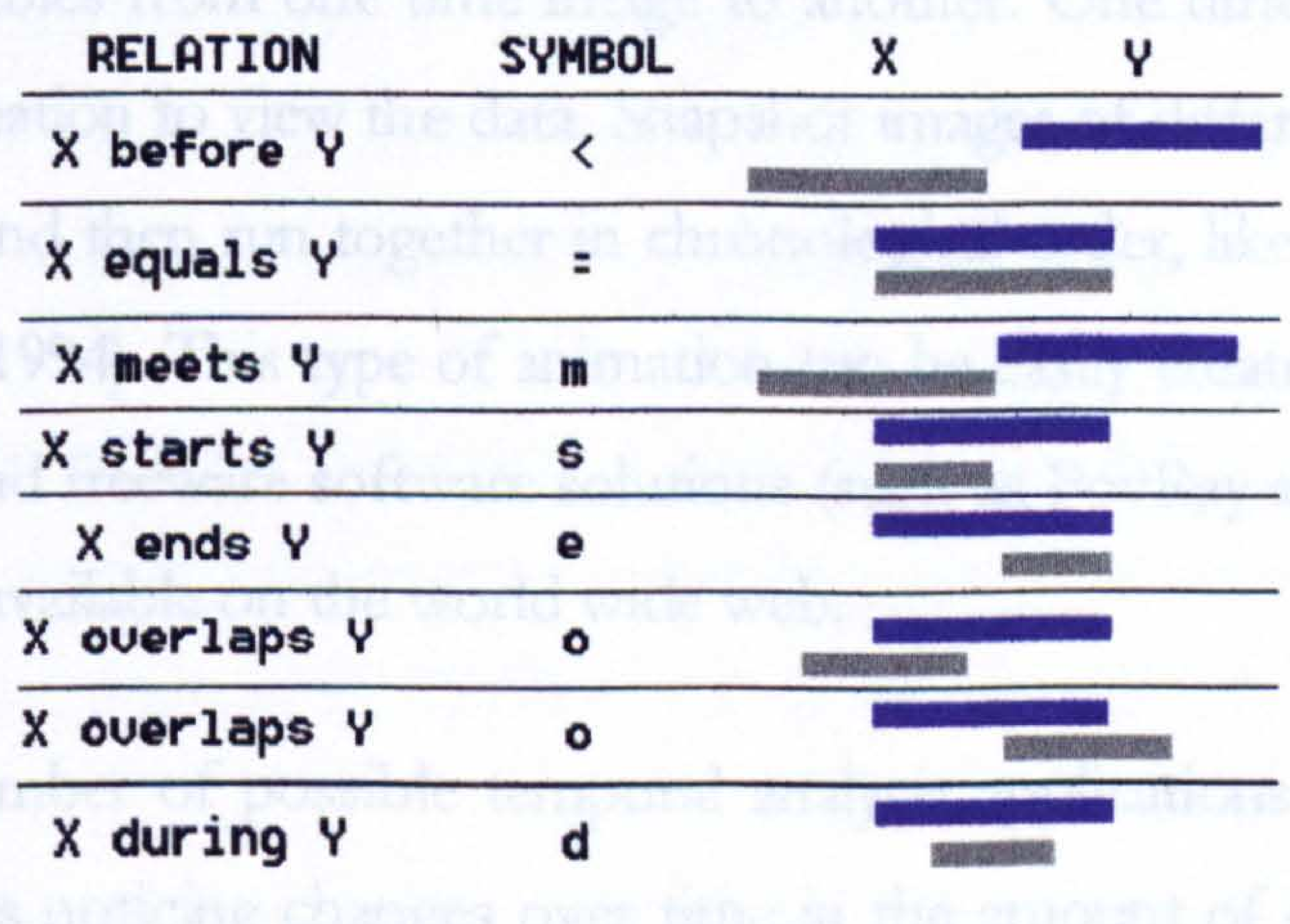


Figure 2-5 Temporal topological relationships.

From Peuquet, 1994, p.455.

Describing spatiotemporal relationships is not easy and the extension to analysis has also created problems. The desire to analyse temporal data has lead to descriptions of particular temporal operators. One of the more comprehensive description of the variety of possible temporal combinations comes from Peuquet (1994) and is applicable if the passage of time can be viewed as a fixed line and events are fixed to this line in some manner. Crime data is rarely definable as one quantum moment in time. A householder might go on holiday to return two weeks later to find they have been the victim of a burglary. They are unable to tie the time of the incident to a specific moment so the crime report records an end time which is two weeks from the possible start time. Therefore when describing crime events, they must be viewed as singularities of variable length along the time line. These types of relationship are shown in Figure 2-5. Once events are described in this manner, Boolean operators can be applied to them, the data becomes easier to handle, and queries can be performed using a standard query language (SQL).

Maps can be used to visualise temporal change in a number of different ways. Traditionally one of two basic methods have been used. The first shows the user a single map depicting the change from a previous state to the end state which is represented by the map, and changes over time are a fairly simple cartographic feature. The second method uses the 'chess map' approach where a number of temporal stages are shown on different maps, allowing the viewer to follow the change in variables from one time image to another. One different approach is the use of animation to view the data. Snapshot images of different 'world states' are generated and then run together in chronological order, like frames in a film (MacEachren, 1994). This type of animation can be easily created and a number of shareware and freeware software solutions (such as PovRay and Dave's Targa Animator) are available on the world wide web.

There are a number of possible temporal analysis applications to police work. There is value is noticing changes over time in the amount of crime in an area. Sudden changes might relate to the release of a well-known criminal, or changes in policing strategy. In a practical policing application, relative temporal change might be more useful than absolute quantities. Practical applications and a conceptual framework for a police spatial and temporal system are presented in chapter 4 (Aoristic crime analysis).

2.4. CRIME PREVENTION AND REPEAT VICTIMISATION

> *Prevention, the [recent Audit Commission] study complains, is unglamorous. It tends to involve giving 'bolts and bars' advice alone. Instead, preventors and investigators should co-operate, the study suggests. They could be making computer assisted analyses of crime data, for example, to see where offences are most likely to occur. (p.30) (Economist, 1994)*

Police crime data is an obvious source that has been used for crime prevention purposes. With limited resources available for crime prevention, coinciding with a lack of bureaucratic enthusiasm for the subject (see quote above), police authorities have resorted to finding ways of targeting their limited crime prevention resources to the most needy and profitable areas. One of the main problems to date has been to make meaningful use of the data available. Identification of the most worthwhile areas may be within the possibilities of the police, but often other authorities have the actual resources:

> *The government may be warming to crime prevention. But its policy is a mess. Responsibility for prevention is shared between central government, local councils and the police. (p.30) (Economist, 1994)*

Vague and inaccurate information is of no use when trying to convince a cash-starved local authority, or central government, to part with money for crime prevention. One of the key areas of crime prevention which has been recognised only in recent years is the area of repeat victimisation. Repeat victimisation is a target area for crime prevention, but as will be seen later there are great difficulties in actually identifying with any accuracy the repeat victims from the mass of police recorded crime data. Most of the related work has been published in criminology journals and for a review of recent quantitative criminology in the UK, see the work by David Farrington (Farrington, 1996).

The benefit to crime prevention of identifying repeat victimisation has been widely recognised (Anderson *et al.*, 1995; Ellingworth *et al.*, 1995; Farrell and Pease, 1993), but the process of distinguishing accurately the repeat locations has always been difficult. Much of the literature concerning the search for, and use of information on repeat victimisation can be found in either criminology journals or Home Office publications. The British Crime Survey is the major source of data for quantitative criminology, and the Manchester Quantitative Criminology Group have carried out a number of surveys based on this data set (Farrington, 1996). Much less research has been based on police crime data. A number of articles highlight the difficulties posed by police data in identification of repeat victims (Anderson *et al.*, 1995; Ellingworth *et al.*, 1995; Farrell and Pease, 1993; Hope, 1995; Read and Oldfield, 1995; Sampson and Phillips, 1995). As already said, the police data tends to be recorded for management information statistics, and not specifically designed for the identification of repeat victimisations.

An example of the difficulties posed by police crime recording systems is provided by David Anderson and his colleagues (Anderson *et al.*, 1995). They attempted to download all of the possible repeat victimisation crime data from the West Yorkshire Crime Information System in ASCII format, and then convert the data for analysis in a Windows statistical package, SPSS. However;

> *Once the data had been downloaded, the process began of manually checking the data and converting it to a form which could be read into an SPSS system file for analysis. The work was massively time consuming. It took four months... (p. 6)*

The majority of repeat victimisations occur within one month of previous (or the first) victimisation (Pease, 1997). This has distinct implications for the location and targeting of crime prevention resources, resources which must be marshalled quickly. However it will be seen that the use of a GIS could rapidly enhance the search facilities and speed up the identification of repeat victims. Once a repeat victim has been identified, crime prevention resources have to be mobilised rapidly to prevent further attacks on the same target.

Crime prevention will undoubtedly be improved if a more efficient means of identifying the repeat victims becomes available. This should be the focus of the available resources. To achieve these goals the police need the means to access the data quickly, certainly within the month time frame for most repeats, and they need the ability to correctly identify the locations where repeats have taken place over the study time period. Until now it has often been this second task which has been the stumbling block for crime prevention strategies.

One potential solution is to take advantage of the spatial nature of georeferenced data. This in now possible as many police forces have access to geographically referenced data. Analysis with GIS has significant advantages over standard databases as 'the key features which differentiate GIS from other information systems are the general focus on spatial entities and relationships' (Maguire, 1991). The inclusion of a geographical reference in the recorded crime data shows potential for new methods of searching for repeat victimisations and gets away from the traditional method of text-based database searches. This possibility is explored in chapter 6 (Identifying repeat victimisation).

2.5. CONCLUSION

This chapter has reviewed critically the main topics as a whole that concern this thesis and identified the aspects where this thesis can make a contribution. The current literature shows that crime mapping work in the United States is more advanced than in the UK, though the emphasis has been on mapping and not on analysis. The identification of crime hotspots, while worthwhile, is yet to develop a satisfactory methodological framework. Chapter 8 (Hotspot analysis) will show this results, at least in part, from a lack of sophistication in the available software.

Literature concerning temporal considerations in crime is rare, demonstrating a lack of enthusiasm for temporal studies of crime. The temporal inexactitude of the commission of crime does not help matters, though the benefits might be valuable to the police service. A conceptual framework for the analysis of this temporally inexact material is necessary before a more thorough temporal analysis can take place.

Finally the practical benefits of a greater comprehension of the mechanisms of repeat victimisation are detailed. The lack of a technical ability to identify quickly repeat victimisation hampers crime prevention efforts, and a more thorough understanding of the social pattern of repeat victimisation would benefit the targeting of scarce resources.

All of these subjects are discussed in more detail in the chapters that follow.

3. Data sources and software

This chapter describes the sources of data used throughout the thesis.

The primary study area of Trent division is described. Trent is the largest of Nottinghamshire's police divisions and has a wide diversity of inner city, suburban and rural regions.

The origins of the police recorded crime data are then discussed, along with the accuracy of the georeferencing that accompanies the crime records. This is discussed in the general context of georeferencing techniques and the use of both the Postcode Address File, available from the Royal Mail, and AddressPoint, available from the Ordnance Survey. The same georeferencing techniques are applied to the newly available police incident data.

Various other data sources, such as the Index of Local Conditions, are introduced and the chapter ends by listing the proprietary software used throughout the thesis.

3.1. THE STUDY AREA

The areas studied in this thesis are all contained within Nottinghamshire Constabulary. Various areal levels of study take place within the thesis. These vary from the force and divisional level to smaller investigations within the boundaries of one sub-divisional station and its beats.

3.1.1. Nottinghamshire Constabulary

Nottinghamshire Constabulary is one of 43 geographical (regional) police forces in the UK. Examples of non-regional forces include the British Transport Police, Atomic Energy Authority Police and the Ministry of Defence Police; all of whom have national coverage. The county is bordered by Leicestershire, Lincolnshire, Derbyshire and South Yorkshire (Figure 3-1).

The force consists of 9 regional divisions (Figure 3-1) responsible for the geographical policing of the county. Each division is controlled by a Chief Superintendent who maintains overall responsibility for all policing operations

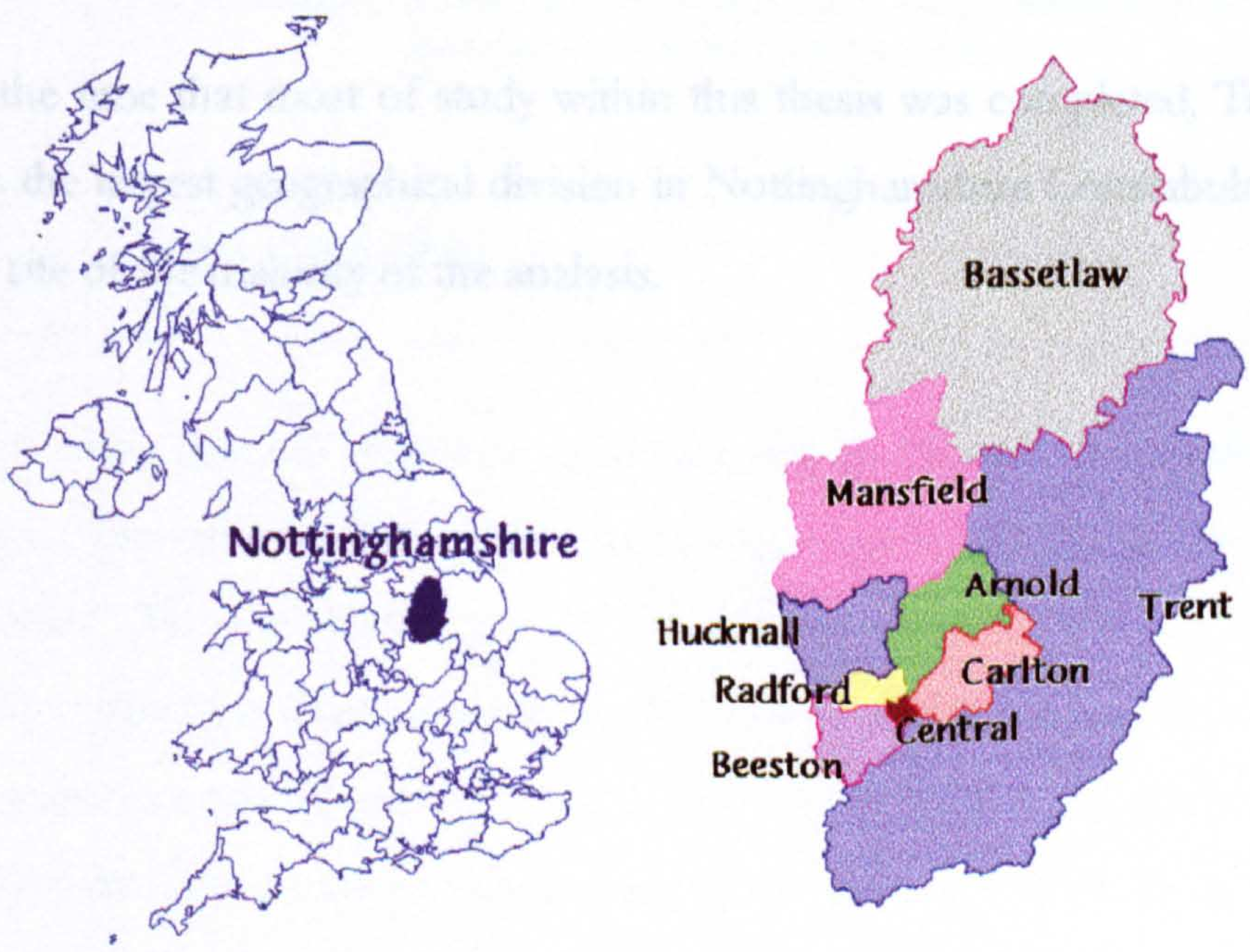


Figure 3-1 Nottinghamshire county and the nine geographical police divisions.

on that area, and who is answerable to the Chief Constable at Force HQ located at Sherwood Lodge, North of Nottingham.

Table 3-1 Nottinghamshire Constabulary rank structure.

Chief Constable
Assistant Chief Constable
Superintendent
Chief Inspector
Inspector
Sergeant
Constable

Although the rank of Chief Superintendent has recently been removed by Nottinghamshire Constabulary, a number of officers remain in this post, including Chief Superintendent Eddie Curtis at Trent division, West Bridgford Police Station.

RANK STRUCTURE

Like the majority of other police forces in the UK, Nottinghamshire Constabulary maintains a standard rank structure and in recent years has discontinued promoting to the rank of Chief Superintendent, although a number of individuals remain in this rank. The force has also dispensed with the rank of commander, which is still in use in the Metropolitan force. The rank structure of Nottinghamshire Constabulary is shown in Table 3-1.

At the time that most of study within this thesis was completed, Trent division was the largest geographical division in Nottinghamshire Constabulary. It is also the site of the majority of the analysis.

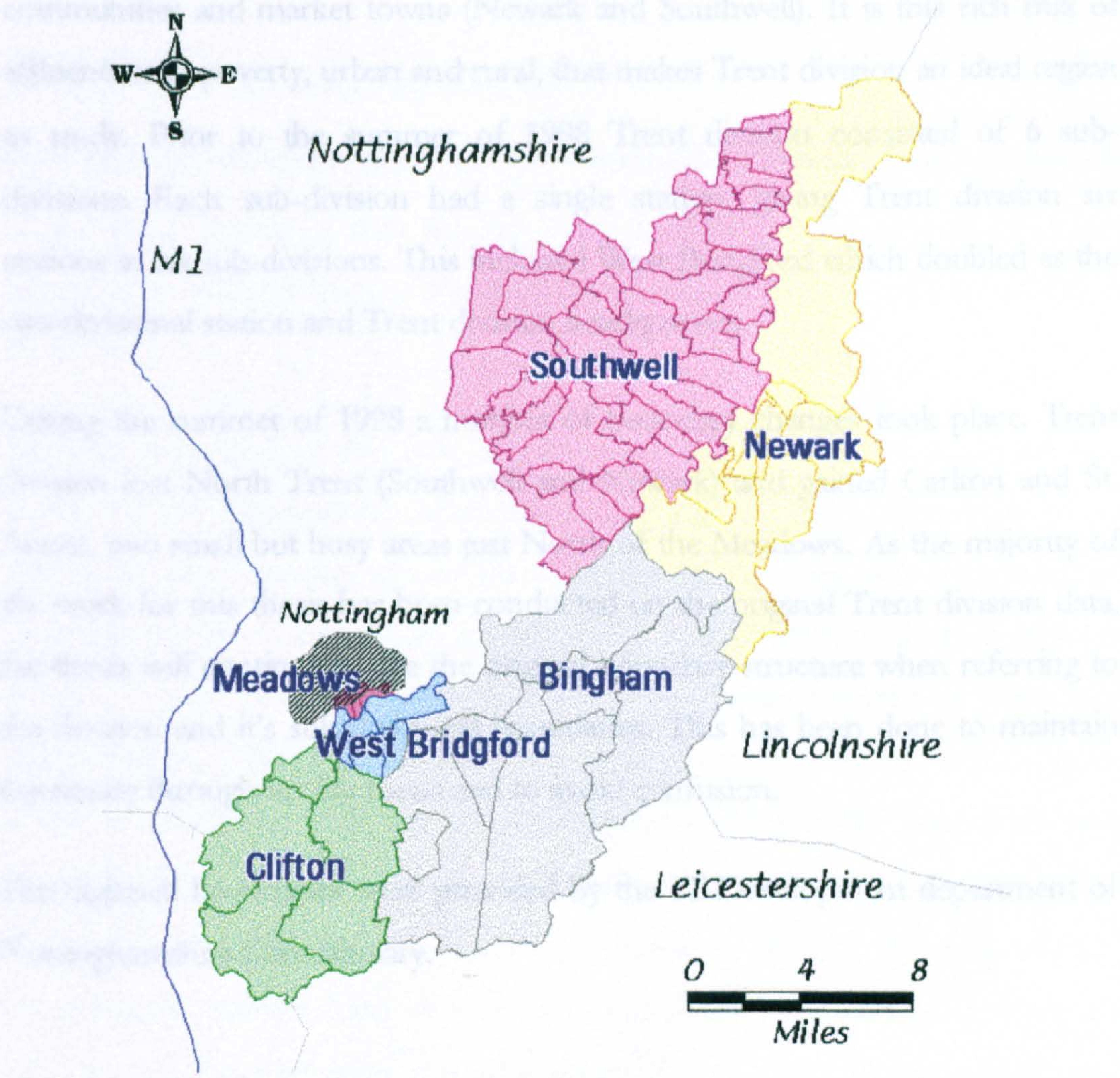


Figure 3-2 The six sub-divisional areas of Trent division, Nottinghamshire Constabulary prior to mid 1998.

The six sub-divisional areas are shown by the blue text. Internal boundaries delimit the individual beats of each sub-divisional station.

3.1.2. Trent division

Trent police division covers most of the South and East of Nottinghamshire and mixes affluent suburbs, council estates and rural villages within one police division. To the West and North it borders other divisions within Nottinghamshire Constabulary, to the East and South are Lincolnshire and Leicestershire (see Figure 3-2), and to the South West is the Eastern corner of Derbyshire. The southern sub-divisions (Meadows and West Bridgford) border the city centre of Nottingham and are a mix of inner city estates and modern suburbia. Further from Nottingham city can be found rural farmland, old mining communities and market towns (Newark and Southwell). It is this rich mix of affluence and poverty, urban and rural, that makes Trent division an ideal region to study. Prior to the summer of 1998 Trent division consisted of 6 sub-divisions. Each sub-division had a single station, giving Trent division six stations in six sub-divisions. This included West Bridgford which doubled as the sub-divisional station and Trent division headquarters.

During the summer of 1998 a number of boundary changes took place. Trent division lost North Trent (Southwell and Newark) and gained Carlton and St. Annes, two small but busy areas just North of the Meadows. As the majority of the work for this thesis has been conducted on the original Trent division data, the thesis will continue to use the original boundary structure when referring to the division and it's sub-divisional boundaries. This has been done to maintain continuity throughout the thesis and to avoid confusion.

The digitised boundaries were provided by the IT Development department of Nottinghamshire Constabulary.

3.2. POLICE CRIME DATA

The primary source of data for this thesis is a two year record of all crimes recorded on Trent division of Nottinghamshire Constabulary from April 1995 to April 1997.

The first download of this type of data was made available for a second year undergraduate dissertation completed by the author. An exabyte tape of all crimes for Trent division was produced by the IT Development department of Nottinghamshire Constabulary in July 1995, and covered a short 3 months worth of data. The dissertation highlighted the future benefits of continued crime analysis of Nottinghamshire police data. Once the main research program began, Nottinghamshire Police were asked (internally by a Chief Superintendent) to repeat the download of data for a longer period, and to include data for the whole force area. Internal political problems within Nottinghamshire Constabulary prevented any data being produced for a number of months. As with the first data provision, this information was provided on an exabyte tape, in pipe (|) delimited ASCII format, and contained a total of 294,797 records. This tape was converted and written to a CD-ROM for ease in handling and manipulation by the author in the department.

The columns of information contained in the ASCII file for each of the nine geographical divisions of Nottinghamshire Constabulary are shown in Appendix A. The number of crime incidents for each division is shown in Figure 3-3.

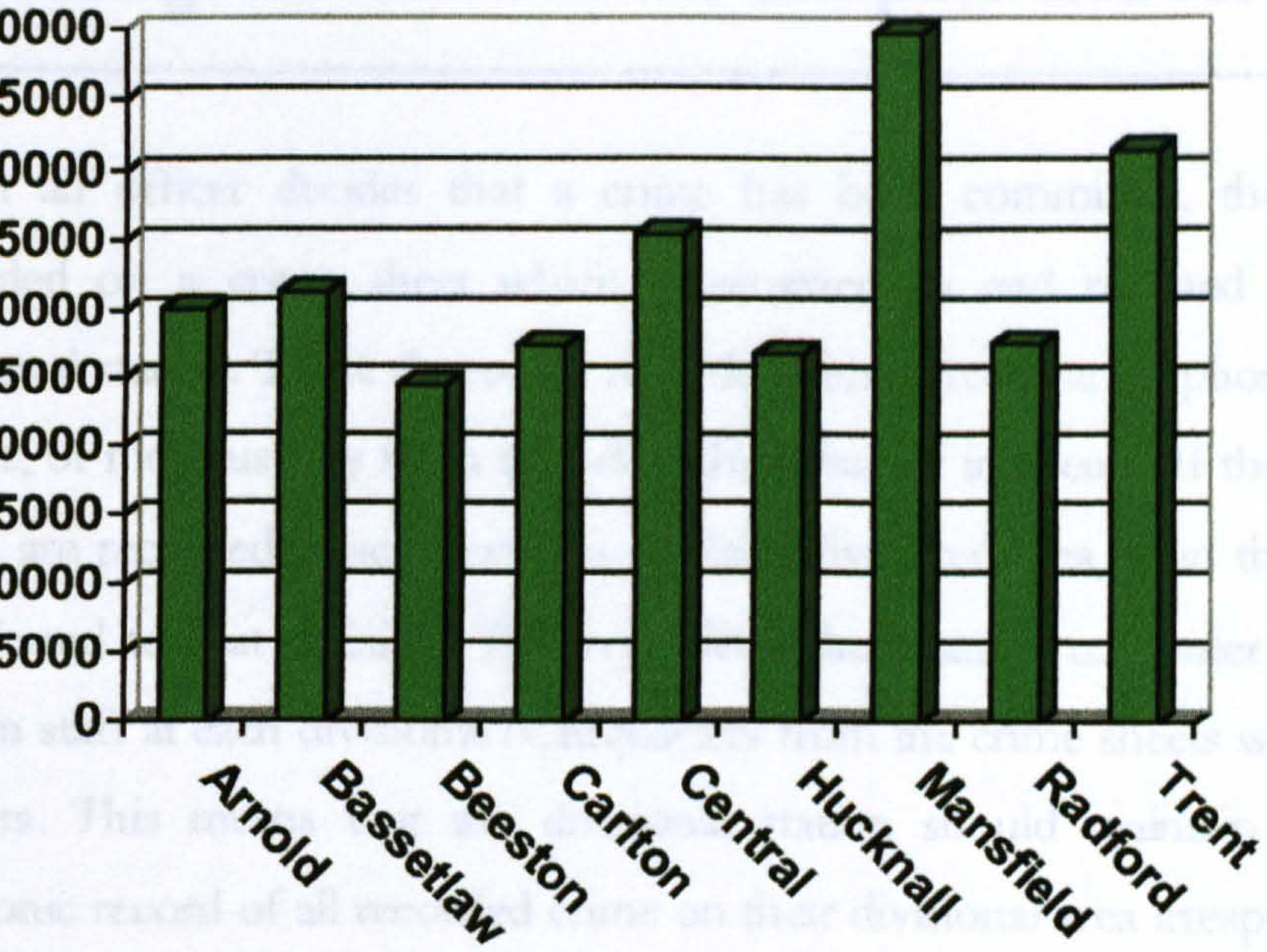


Figure 3-3 Reported crime in Nottinghamshire by Police division (April 1995 to April 1997)

The grid location provided in column 16 (variable GRIDCODE in Appendix A) is a single number made up of the 0.1 metre resolution Ordnance Survey (OS) Easting co-ordinate immediately followed by the 0.1 metre Northing. New columns were added to split this co-ordinate down into its constituent parts. For example; 045681153445248 became Easting 456811.5, and Northing 344524.8. This was necessary as the next stage took the file into MapInfo and converted the text data into a GIS-type mappable table.

Tables of data such as the Nottinghamshire Police crime data can be made mappable in a number of ways. Area data can be mapped to existing polygons, however location values (such as the crime data) can be converted into point-mappable data by setting a co-ordinate system such as British National Grid and creating points (using the MapInfo *Create Points* command) based on an x and y co-ordinate column. The Easting and Northing derived variables were used for these columns.

3.2.1. Original source of the computerised records

When an officer decides that a crime has been committed, the details are recorded on a crime sheet which is returned to and retained at the local divisional station. These sheets are recorded either over the telephone for minor crimes, or more usually from attending the scene of incidents. If the details of a crime are recorded which occur in another divisional area, then the details are transferred to that division. This recorded crime data is computer encoded by civilian staff at each divisional headquarters from the crime sheets written by the officers. This means that the divisional station should maintain a complete electronic record of all recorded crime on their divisional area irrespective of the point of recording.

Various criteria are applied to the recording methods. Generally one crime generates one crime sheet, however there are circumstances when this is changed. For example, when a number of events occur in one incident then they may be recorded on one crime sheet. This could happen when criminal damage occurs at the scene of a burglary. Although different crimes, the criminal damage would be recorded as additional information on the burglary crime sheet. A continuous repeat of a similar incident would also generate a single crime sheet. This can occur in cases of repeated sexual assault which have been unreported for a long time, where a large number of separate incidents have occurred over a number of years. In this example a single specimen crime sheet will be used to record the pattern of offences.

Aspects of the various categories should be briefly explained at this point. Variables HOMAJOR, HOMINOR, FORCEMAJOR, FORCEMINOR (columns 4 to 7 in Appendix A) are different interpretations of the same incident. The civilian operator at the divisional station suggests a (Nottinghamshire Police) force major (FORCEMAJOR) and minor (FORCEMINOR) classification based on the information available. An example of this might be a major classification of BURGLARY and a minor classification of AGGRAVATED (involving physical violence). These are used as general classifications within the force to permit simple queries of the data using Standard Query Language (SQL) packages such as those available in Microsoft Access and Excel. The HOMAJOR (Home Office major crime) and HOminor (Home Office minor crime) categories are added by the force headquarters based at Sherwood Lodge. These are the official Home Office categories which are used nationally for statistical monitoring of crime rates. These are added by the specially trained staff at the force HQ who are presented with a wider variety of possible classifications. ADDCODE, STREETCODE and DISTCODE (columns 14 to 16) are unique local variables for street names. They were part of an original attempt by Nottinghamshire Constabulary to develop a local gazetteer of Nottinghamshire streets and locations, and were to be used as an additional SQL-based variable. Each address or street was given a numerical code as an identifier on their system, and as new streets or locations became known, they were added to the system and given the next available number. This has been superseded by the use of properly georeferenced co-ordinates based on the British National Grid (see next section). The local street reference system has at present no application within the force.

Columns 19 to 24 relate to details of the property involved in the crime. This part of the crime sheet is only usually completed in the case of burglary. The civilian operator interprets the description of the burglary from the officers written notes on the crime sheet and from this decides on categories for the type of premises (PREMCODE e.g. ground floor flat, bakery shop), the method of entry (MOECODE e.g. breaking window with implement) and the point of entry (POECODE e.g. rear ground floor window).

PROPSTOLEN and PROPDAMAGED (columns 25 and 26) are the estimate of the value (in sterling) of the property stolen (PROPSTOLEN) or damaged (PROPDAMAGED) in the incident. This is generally based on the estimation of the property worth from the victim. Ethnicity, gender and age details of the victim are recorded (columns 27, 28 and 29) which can aid in the selective monitoring of calls such as racial incidents, and crimes near schools.

Vehicle details (columns 30 to 37) are recorded not just for stolen vehicles, but also for vehicles which have not been stolen but have property stolen from them or are damaged. These details are also recorded for vehicles used in crime, such as getaway cars and vehicles used in ram-raiding.

Finally it should be noted that the dates and times of crimes are rarely known with any precision. If a burglary happens when a family are on holiday (for example), the time of the incident may not be known more accurately than to a two week period. In the majority of cases such as these, the incident is recorded as having a FROMTIME, FROMDATE, TOTIME, and TODATE (columns 8 to 11). When an incident time is known exactly the FROM variables are the only ones completed. The use of this temporal inaccuracy for analysis purposes is explored in the next chapter.

The recorded crime data relates to all crimes reported to the police and classified as such by the police. The under-reporting of crime to the police has been discussed in the previous chapter, as has the effect of police officer interpretation and classification of incidents as either crime or non-crime.

3.2.2. Georeferencing

In the last few years, a number of national police forces (including Nottinghamshire) have been georeferencing their crime incident records by including the National Grid co-ordinates of the event on their systems. In the case of Nottinghamshire, they began by tying in the location of the crime incident to the Postcode Address File (PAF) which has a spatial resolution of 100 metres for the full Postcode centroid. A number of authors comment on the accuracy and usefulness (or not) of the PAF (Gatrell, 1989; Raper *et al.*, 1992). Raper showed that while a spatial error of around 60-70 metres was the average for a study area in Cumbria, there were more problems if the 100 metre postcodes were plotted within the boundary of their respective enumeration district. When the postcode centroids and enumeration district boundaries were directly compared, often the centroid fell outside the boundary for the enumeration district to which the postcode region belonged. In this study we are more concerned with the level of inaccuracy between the location of the property and the 100 metre generalisation. Recently (since 1994) the

Nottinghamshire Crime Recording Interim System (CRIS) has been updated to take advantage of Address-Point data, a commercial package available from the Ordnance Survey (OS) capable of identifying each property in the country with a unique 0.1 metre resolution National Grid reference. This data source is updated biannually. The CRIS upgrade has coincided with a more rigid data entry dialog on CRIS, which forces the operator to find an address which is known to exist. This is designed to prevent the problem of having various address fields for the same location. As long as the central database of properties, addresses and georeferences in the county is maintained, there should be few locations which are not known to the system. The CRIS system automatically assigns the grid location which corresponds to the address or location chosen by the operator.

Whilst the system will eventually contain only Address-Point grid references, a number of georeferenced locations derived from the PAF still exist in the gazetteer. The PAF agglomerates all addresses which share a postcode and gives them the same 100 metre resolution grid reference. Postcode grid references which are very close together can share grid references. This can mean that a number of different addresses can have the same postcode. Address-Point data is a considerable improvement. However, while the Nottinghamshire system uses Ordnance Survey and Royal Mail products, it is not always necessary to use the National Grid co-ordinates. Hirschfield and colleagues (Hirschfield *et al.*, 1995) at Liverpool University have been analysing crime incident data using the Merseyside Address Referencing System (MARS), which forms the basis of the Command and Control and Crime Incident Reporting System used by Merseyside Police. This database was available to those researchers, as was that part of MARS which provided a one-metre resolution grid reference for each property in the study area. All premises are identified by Unique Property Reference Numbers (UPRN). The UPRN is recorded on every crime report and in their command and control data capture operation. Although the UPRN is unique, the grid references can be shared by different addresses.

The MARS system allocates a one metre resolution grid reference even though a high resolution OS grid reference is not sufficient to deal with the address problem of multi-occupancy buildings. A future solution to this problem might be the use of Ordnance Survey Address-Point Reference (OSAPR) which identifies different addresses within the same building. This is a slightly different approach to the one adopted by Nottinghamshire Constabulary. Without the benefit of inheriting a local address referencing system such as MARS, Nottinghamshire Constabulary have created their own gazetteer of addresses within the county, to which an Address-Point grid location is attached. This is a accurate to 0.1 metres, though suffers the same problem of identifying individual elements of multi-occupancy buildings.

3.2.3. Distribution of crime by day and time

The following section briefly outlines the overall distribution of crime in Nottinghamshire. The distribution of crime across Nottinghamshire is fairly uniform though Mansfield and Trent division receive the highest levels of recorded crime (Figure 3-4 below and Figure 3-3 on page 50). Daily distribution of all crime is also even with a slight peak on Tuesdays.

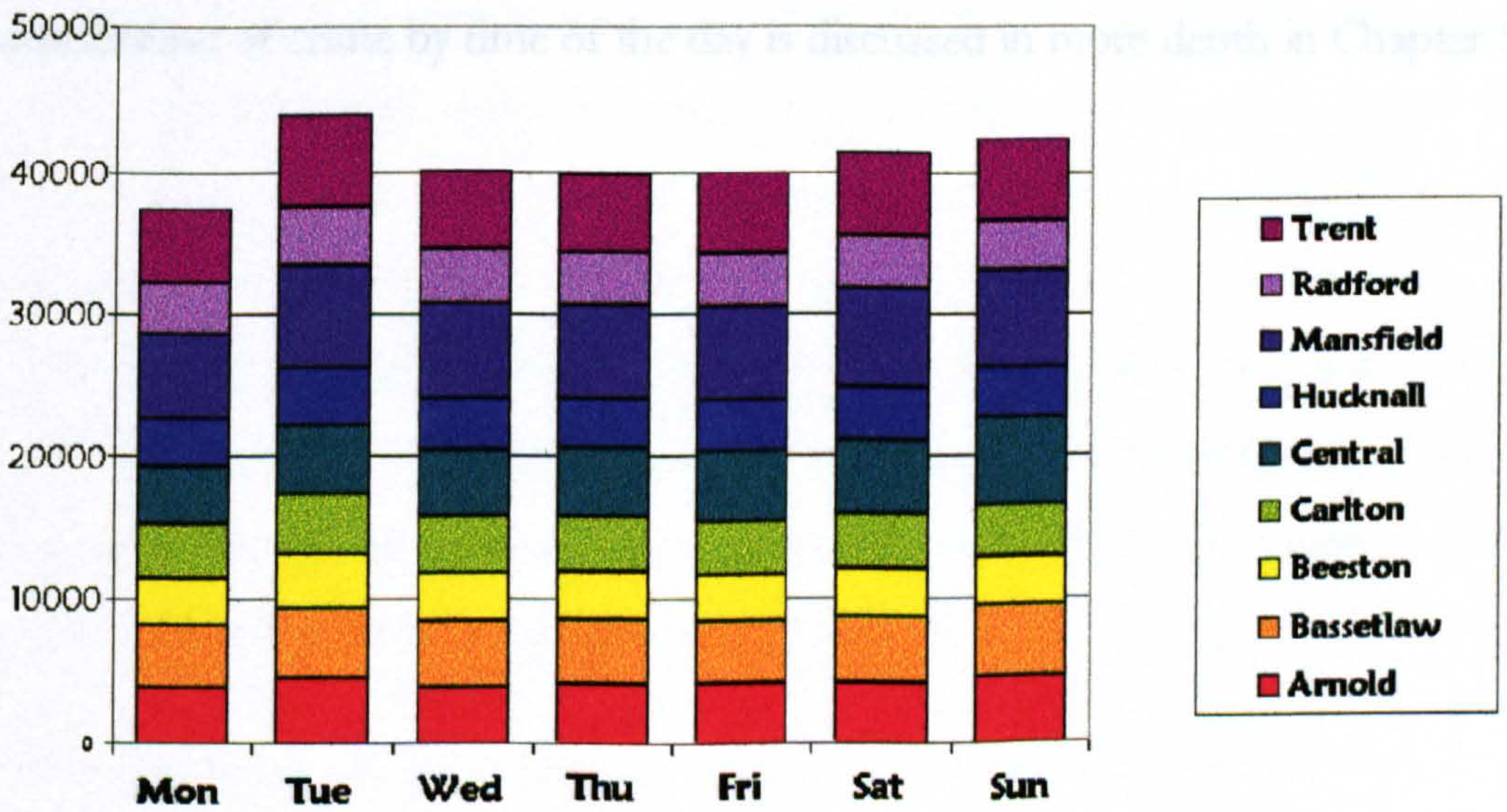


Figure 3-4 Crime in Nottinghamshire's police divisions by day of the week (April 95 – April 97).

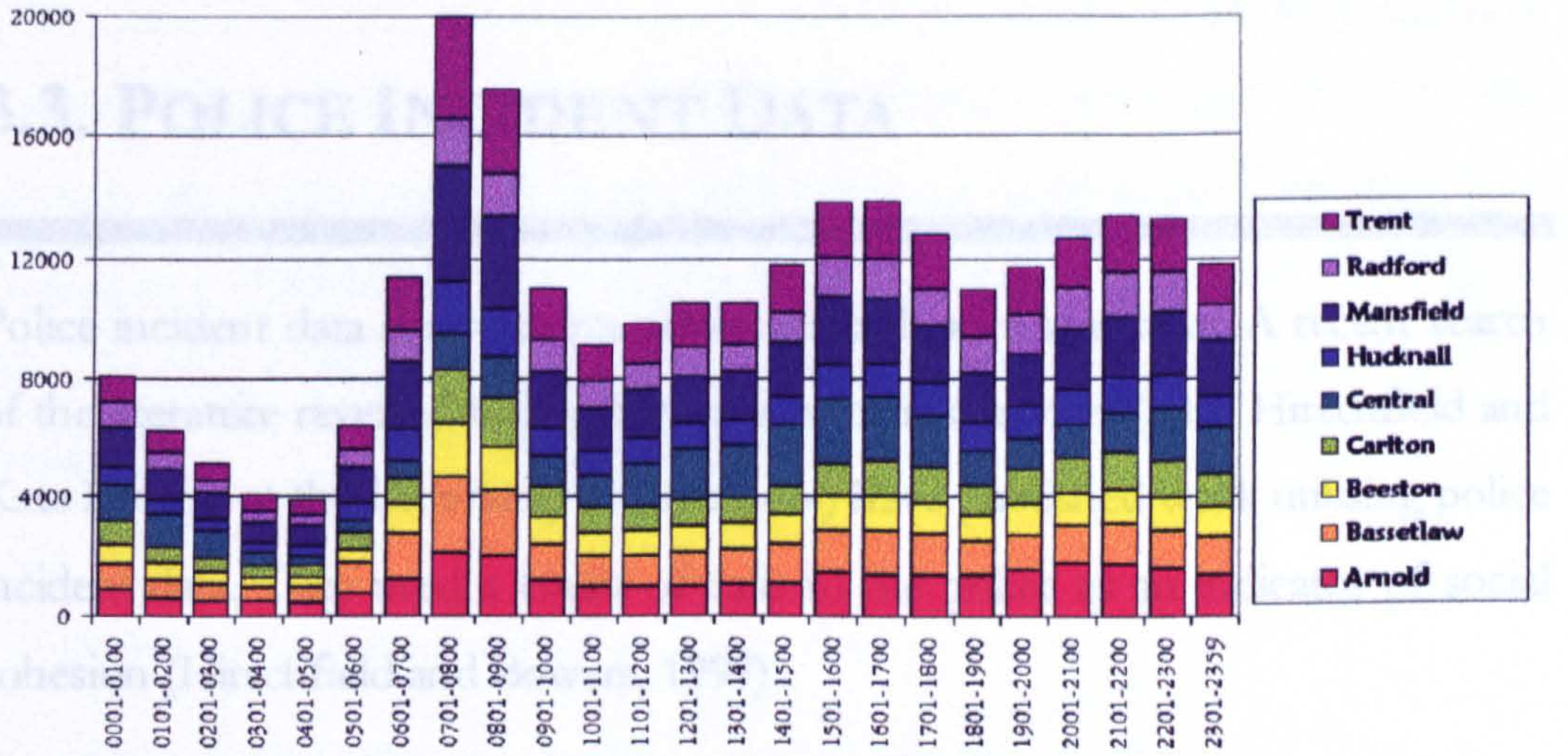


Figure 3-5 Crime in Nottinghamshire's police divisions by time of incident.

The time of incident is taken here from the TOtime field in the crime data.

The distribution of crimes across the county by time is shown in Figure 3-5. There is a significant rise in the number of crimes which show a TOtime between 7am and 9am and it is likely that this is due to the number of overnight offences (such as night-time burglaries) committed at commercial premises which are discovered and reported when workers arrive in the morning. The distribution of crime by time of the day is discussed in more depth in Chapter 5.

3.3. POLICE INCIDENT DATA

Police incident data is a resource rarely available to researchers. A recent search of the literature revealed that only one team of researchers (Alex Hirschfield and Kate Bowers at the University of Liverpool) have published work utilising police incident data. They used a count of calls to the police as an indicator of social cohesion (Hirschfield and Bowers, 1997).

Police incident data is a collection of records relating to incidents brought to the attention of the police. These incidents are not crimes *per se*, though many incidents end up being repeated in the recorded crime data records. An 'incident' can be created in a number of ways. One of the most common is when a member of the public telephones the police to report a crime. This phone call can be either to a central 999 system or direct to the station. The details are taken by the telephone operator and are typed directly into an incident log created on the operator's computer screen. Each incident is given a unique reference number. Once an incident has been recorded the incident is passed automatically to the relevant station that then assign an officer to attend and investigate. The incident log records the number of the attending officer, when they were assigned and the eventual outcome of the call. The range of outcomes can include the crime reference number (if a crime is recorded) or a text field to record other forms of result. An example of the latter would be the recording of 'no cause for police action' against an incident where the police found subsequently they were not required to attend.

Other incidents can be created as a record of community meetings which an officer would be required to attend, or can be created by officers themselves. For example a police officer calling for assistance on his personal radio would cause the generation of an incident. The police incident log can be seen as a record of the daily activity of the police officers on an area.

These records have only recently been recorded on computer in any force. When this author was a constable at a police station in East London during the mid

1980's incidents were recorded on pieces of paper and kept in a specially designed folder. Calls from the public were scribbled down and added to the message pad. Calls recorded by the 999 operators at Scotland Yard were sent through by a crude telex system and were printed out in a cupboard. Any urgent 999 calls rang a bell in the cupboard at the same time to attract the operator's attention!

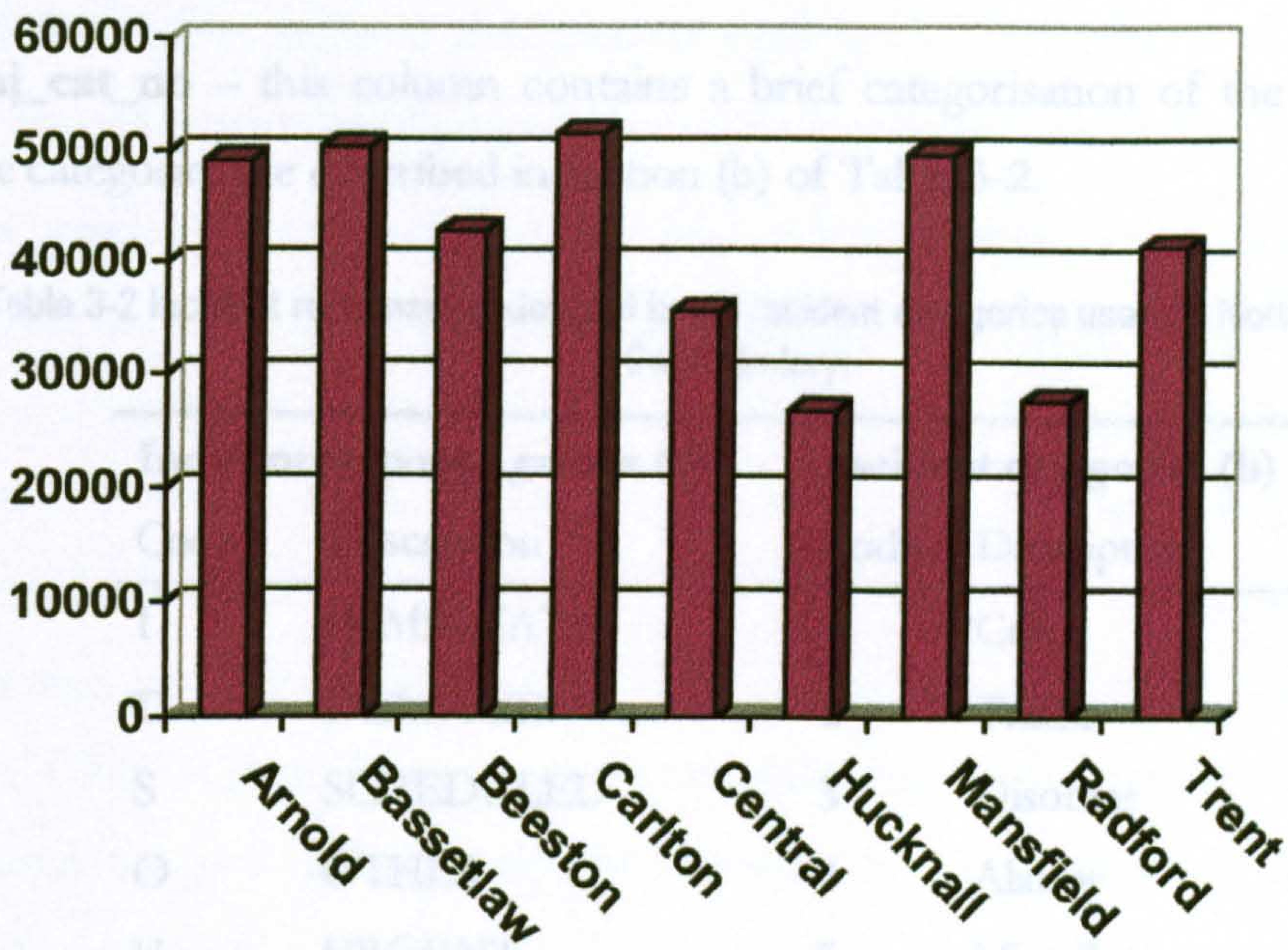


Figure 3-6 Volume of Nottinghamshire Constabulary incidents recorded by divisions (January 1997 to April 1998)

The data made available by Nottinghamshire Constabulary was a limited subset of the full data, containing 375,180 records. The full listing of available data is in Appendix B and the volume of incident data on each division is shown in Figure 3-6. Of particular interest are the following columns:

Global_Inc_No – this is a unique reference number for Nottinghamshire Constabulary.

Incident_Grade – this code is the response grading attached by the officer who initially received the call. The officer makes a judgement as to the urgency of the call and allocates a code as appropriate.

Table 3-2 describes the range of options for both incident grade (a) and incident categories (b).

Gridref – the OS National Grid reference is attached to each record. The process is the same as described earlier for recorded crime data.

Initial_result – the initial result code as decided by the officer who attends the call.

Maj_cat_no – this column contains a brief categorisation of the type of call. The categories are described in section (b) of Table 3-2.

Table 3-2 Incident response grades and broad incident categories used by Nottinghamshire Constabulary.

Incident response grades (a)		**Incident categories (b)**	
Code	Description	Code	Description
I	IMMEDIATE	1	Crime
D	DELAYED	2	Traffic
S	SCHEDULED	3	Disorder
O	OTHER	4	Alarms
U	URGENT	5	Miscellaneous
		6	Non Incidents

There are five grades of call in Nottinghamshire Constabulary (Table 3-2a), which are defined as follows (source: PC Steve Medcalf, IT supervisor, Trent division, Nottinghamshire Constabulary):

Immediate - Used in life-threatening cases or where an offence is in progress. This should be responded to within 10 minutes in urban areas and 15 minutes in rural areas.

Urgent - Where vulnerable persons are in distress, where there is imminent threat to property, where a burglar alarm is activated or where a burglary has been committed. This is in order to preserve evidence at a scene which may be lost. Response time should be within 30 minutes in all areas.

Delayed - All other Incidents. These should be responded to within 24 hours.

Scheduled - Where a specific appointment has been made. Usually within 72 hours.

Other - Answered and dealt with over the phone where no police attendance is required.

3.3.1. Difficulties encountered with incident data

One application of incident data explored later in the thesis is the interaction between incidents and recorded crime. Incidents are created when there is a need to record officially an action on a central database. An incident will be created when a member of the public contacts the police to report a burglary. The incident will record the number of the officer, when they arrived and the crime reference number of the burglary report. The burglary report will be entered onto CRIS (Crime Recording Interim System). Unfortunately the only link between the two different computer systems is a column in the incident data which contains the unique crime record number if a formal record of a crime has been made. For some random reason this column of incident data was not included.

Attempting to extract the duplicate records is problematic due to the nature of recording of both incidents and crimes. If a member of the public calls 999 to report a crime then an incident will be constructed. However if a member of the public attends the front counter of a police station and the crime is recordable at the station, an incident will not be made if there is no need for further immediate police involvement. Only if an officer is required to attend the scene will an incident report have to be made. Because there is no one-to-one relationship between crimes and incidents it is difficult to accurately identify which incidents have generated crime reports, or which crime reports have originated from recorded incidents. The CRIS crime information does not record the incident log that initiated the police response (if any).

3.3.2. Incidents by day and time

Although the emphasis in this thesis is the analysis which is possible at a divisional and sub-divisional level, the following section briefly outlines the distribution of recorded incidents by day of the week and time.

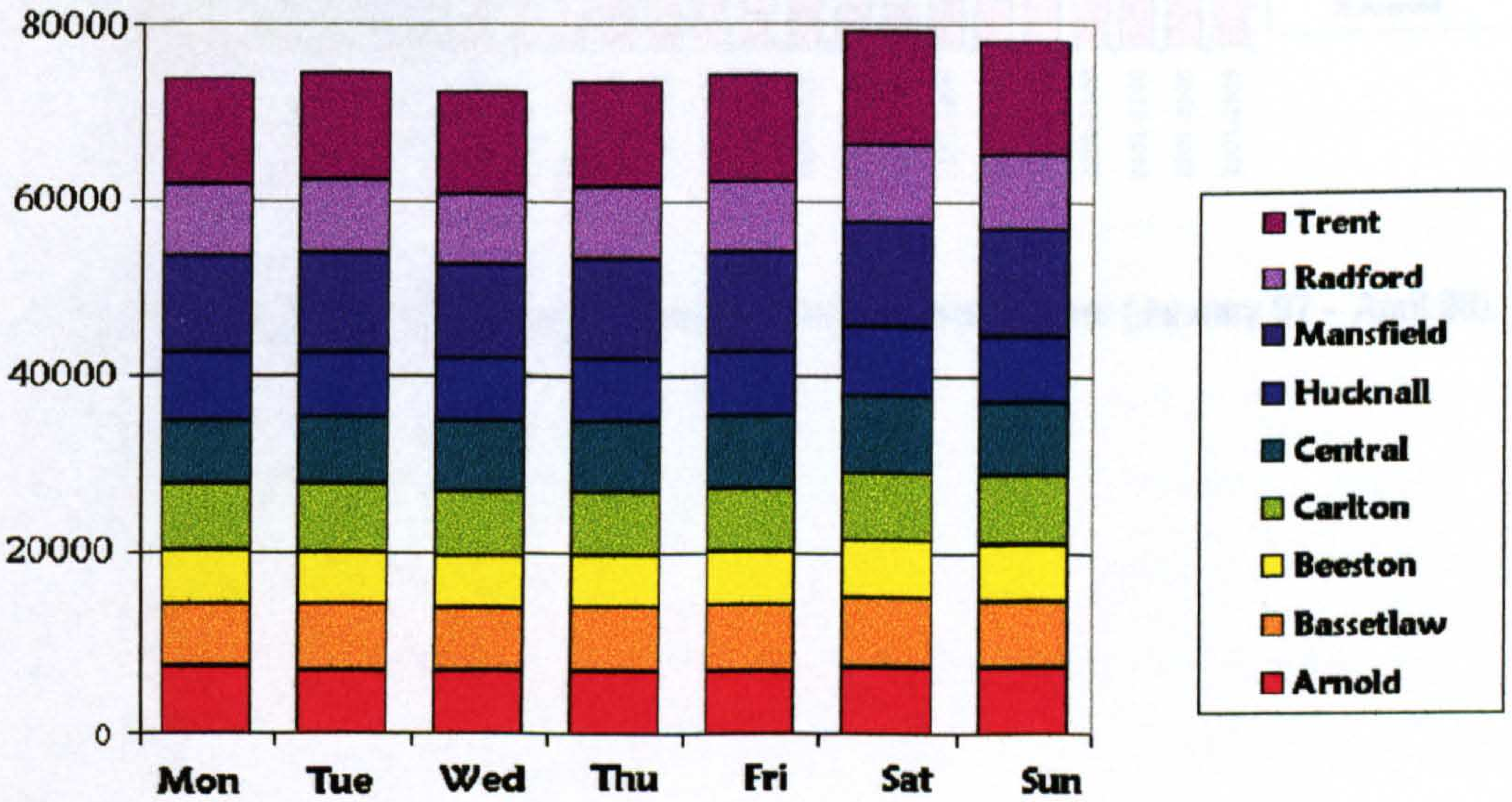


Figure 3-7 Incidents in Nottinghamshire's police divisions by day of the week (January 97 – April 98).

Figure 3-7 shows the distribution of incidents across the nine geographical divisions of Nottinghamshire Police for the period January 1997 to April 1998. As with the crime data, the busiest areas are Mansfield and Trent divisions. There is little variation in the number of incidents recorded on each day.

A histogram on incidents by time of the day (Figure 3-8) shows that requirement for police assistance grows steadily from early morning reaching a peak in the early evening (8pm-9pm) before tailing off to a lower number of calls for service in the hours after midnight.

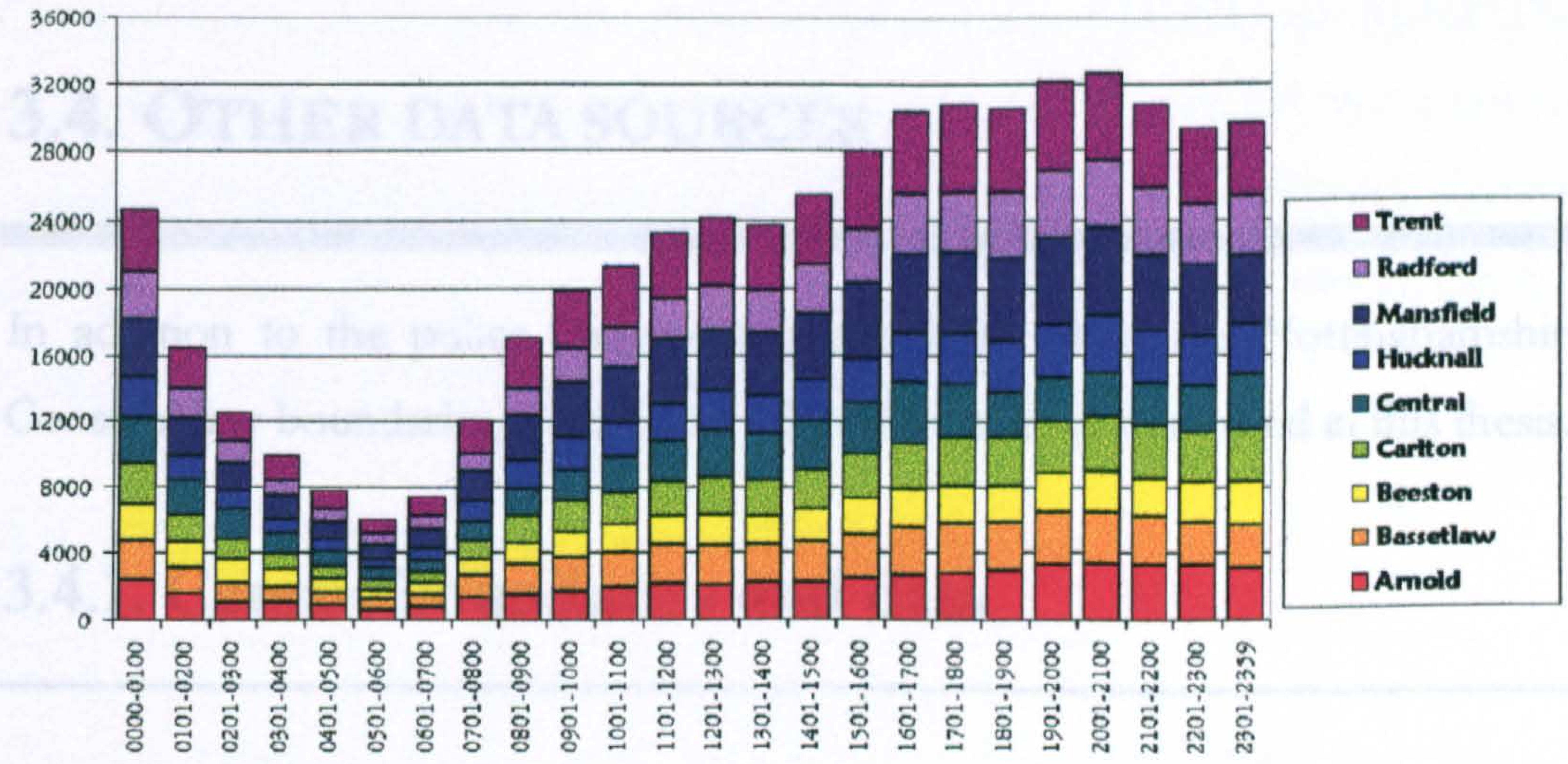


Figure 3-8 Incidents in Nottinghamshire's police divisions by time (January 97 – April 98).

3.4. OTHER DATA SOURCES

In addition to the police crime data, georeferenced to the Nottinghamshire Constabulary boundaries, a variety of other data sources were used in this thesis.

3.4.1. Census boundaries and data

Census variables are taken from the 1991 population census of England. The data used in Chapter 5 (Burglary, victimisation and social deprivation) relating to total numbers of populated houses in each enumeration district is taken from the comma delimited files which accompany the Map91 software. Map91 is a mapping product which enables simple cartographic representation of census information. The enumeration district boundaries are taken from the data available from the Manchester University Data Service (MIDAS) through the

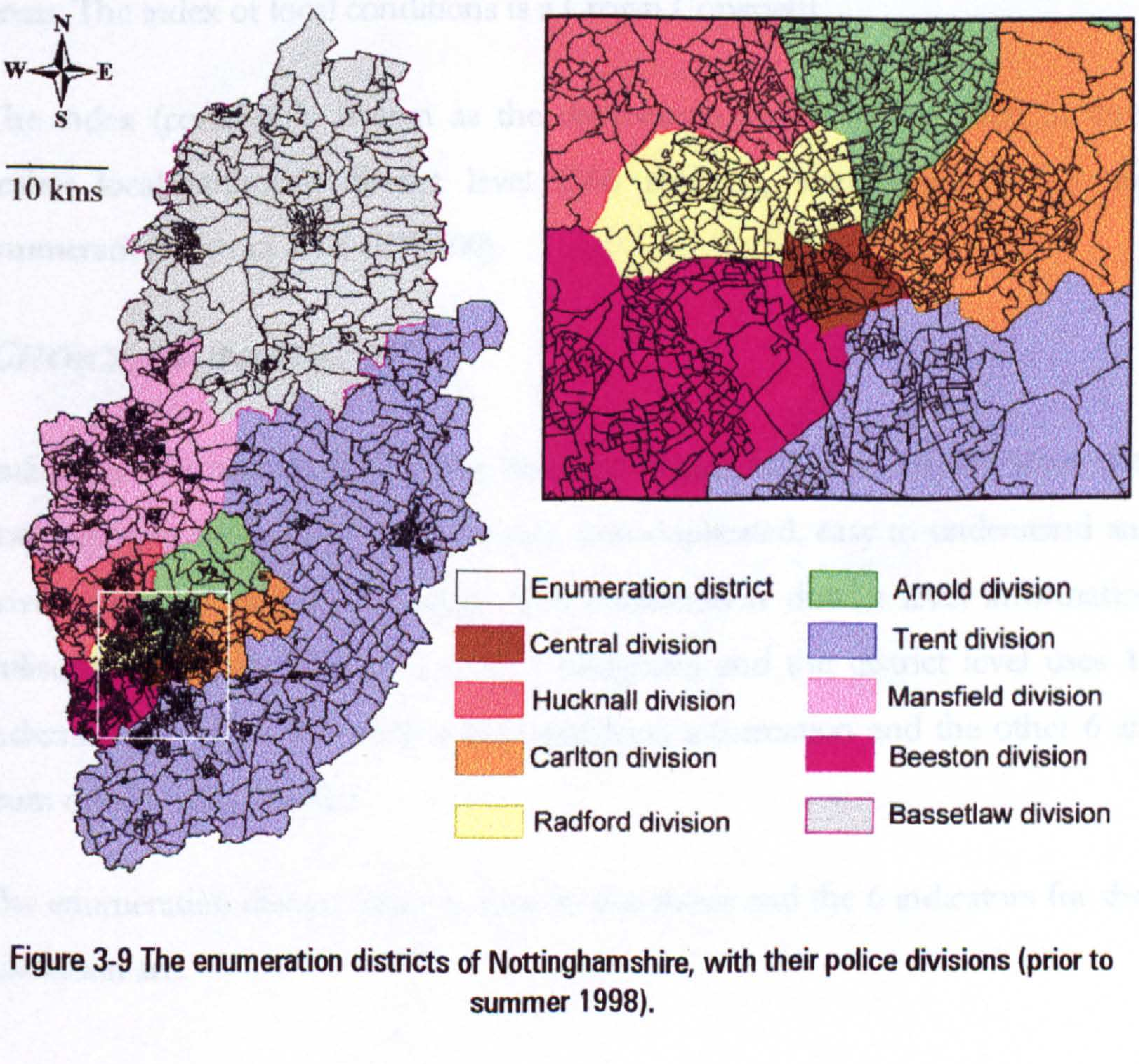


Figure 3-9 The enumeration districts of Nottinghamshire, with their police divisions (prior to summer 1998).

Combined National Higher Education Software Team (CHEST) agreement. The enumeration districts of Nottinghamshire are shown in Figure 3-9 with a colour overlay of the divisional boundaries of Nottinghamshire Constabulary. Central division in the centre of the inset map contains the Central Business District (CBD) and is the city centre of Nottingham.

3.4.2. The Index of Local Conditions

The Department of the Environment's Index of Local Conditions is a measure of relative levels of deprivation across all areas of England, based on the results from the 1991 Census of Population (Environment, 1995). The method of calculation originated in the z-scores calculated from the 1981 census and was computed by the Centre for Urban Policy Studies at the University of Manchester. The revised 1991 index is designed to encompass a number of indicators covering economic, social, housing and other issues, and combine them into an overall measure of deprivation. This countywide analysis allows comparisons to be made and permits the creation of a league table of ranked areas. The index of local conditions is a Crown Copyright.

The index (commonly known as the deprivation index) is calculated at three scales; local authority district level (366 regions), ward level (8,600) and enumeration district level (101,000).

CHOICE OF INDICATORS

Indicators which contribute to the final deprivation index are chosen from data considered to be from robust data sets, non-duplicated, easy to understand and cover the main deprivation areas. The enumeration district level information utilises 6 indicators, the ward level 7 categories and the district level uses 13 indicators, of which 7 are from the ward level information and the other 6 are from outside data sources.

The enumeration district index is used in this thesis and the 6 indicators for that resolution are:

- unemployment,
- children in low earning households,
- overcrowded housing,
- housing lacking basic amenities,
- households with no car,
- children in 'unsuitable' accommodation.

A more detailed description of these variables is available elsewhere (Environment, 1995), where the methodology for creating the index is described. A score of zero at the national level indicates the norm for England. A positive score shows a relatively high level of deprivation and a negative score indicates that the region has a relative level of affluence. Examples at the district level demonstrate the range of values across the UK: Newham and Tower Hamlets in

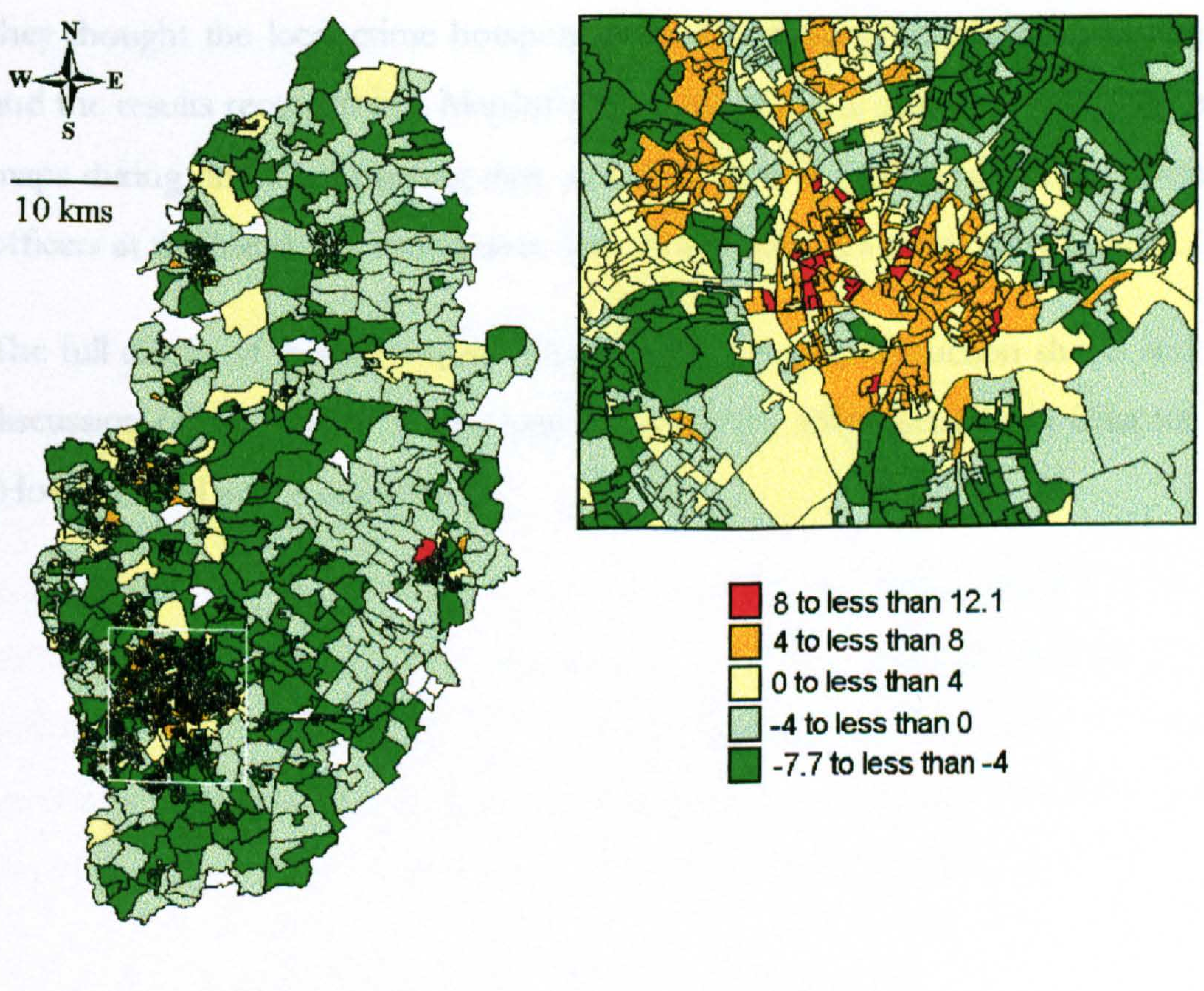


Figure 3-10 1991 Deprivation index values for Nottinghamshire.

Note that a positive value indicates greater relative deprivation.

the deprived East End of London have district deprivation values of over 35, Slough and Hereford exhibit values close to the national mean (0.0) and districts such as Stratford-on-Avon (-25.35) and the Cotswolds (-26.03) are amongst the most affluent (negatively deprived) regions in the country.

The scores for Nottinghamshire are shown in Figure 3-10 where it can be seen that, in keeping with most urban models, the city centre of Nottingham displays greater relative deprivation than the rural and suburban areas.

3.4.3. Qualitative survey data

Survey data collected in the field by the author is used in a study of police perception of crime hotspots. The survey was conducted at three sub-divisional stations on Trent division; West Bridgford, Clifton and the Meadows, during February and March 1998. Operational officers (those concerned with actually policing the streets of Trent division) were interviewed and presented with a number of maps of their beats. They were asked to mark on the maps where they thought the local crime hotspots were. The returned maps were digitised and the results recorded in a MapInfo table. Most of the officers were given the maps during divisional training days and were given an identical briefing. Some officers at the Meadows station were seen individually in a pilot of the project.

The full details of this survey, numbers involved, maps, instruction sheets and a discussion of the validity of this type of survey are fully discussed in Chapter 9 (Hotspots and police perception).

3.5. SOFTWARE

A number of programs were used for this thesis. The main analyses were conducted using MapInfo, Microsoft Excel and SPSS. In addition to this a number of customised programs were written in MapBasic and Visual C++. The main programs used in the thesis are as follows:

- Adobe Premiere - Animation creation.
- MapInfo (versions 4.1, 4.5 and 5) - GIS, mapping, analysis.
- Microsoft Excel - analysis.
- SPSS for Windows (version 8.0) - statistical analysis.
- Microsoft Access - selection of incident data.
- WordStar (version 6) - handling of large ASCII data files.
- MapBasic (version 4.5) - programming of temporal and spatial queries.
- Visual C++ (version 4.2) – programming of temporal and spatial queries.

3.5.1. Programming

The nature of this GIS-based thesis required the customisation of various programs to enhance and speed up the analytical process. Throughout the study period macros were employed within Microsoft Access and Microsoft Excel to perform various analyses. For more specific applications, Visual C++ was used to create unique programs from scratch by the author, and MapBasic was employed to customise analysis within the GIS package, MapInfo.

VISUAL C++

Visual C++ is a powerful tool for building stand alone 32-bit applications for Windows 95, and a number of the programs written for this thesis are written in Visual C++, using the Microsoft Foundation Class library. This library permits functions and classes, written by other developers, to be used within applications developed using the Visual C++ development environment. These are integrated within the applications as fully object-orientated programs.

Visual C++ was used in preference over an available DOS-based Turbo C++ compiler. Although Turbo C is a procedural method of programming C and therefore easier to program, it is not capable of producing fully functioning Windows programs and also has memory limitations.

The current release of Microsoft Visual C++ is version 6. The programs written for this thesis used version 4.2 with MFC version 2.0.

MAPBASIC

MapBasic is the programming language which allows a user to customise and automate MapInfo. It comprises a language of over 300 functions and statements, and a development environment with a text editor, compiler and a linking package to allow various modules in a large program to be separately written.

The language structure and appearance is not dissimilar to modern Basic type languages, such as Visual Basic. Visual Basic is able to generate stand alone programs, but MapBasic compiled programs to run within MapInfo. Because of this, programs are able to utilise the geographical aspects of MapInfo and manipulate the geographical data-management capabilities of MapInfo tables.

The current release of MapBasic is version 5.0.

4. Aoristic crime analysis

This chapter investigates the temporal accuracy problem in crime records and presents a conceptual framework for the temporal analysis of aoristic (temporally unspecific) crime data. This analysis monitors the change in crime patterns over time and can be applied to arbitrary regions, and especially areas smaller than police beat boundaries which have been traditionally the smallest resolution areal unit studied within most police forces. It can also focus the examination on lower crime areas, which can be overlooked in other forms of analysis. Crime data often lacks temporal definition and three different methods of temporal search technique are compared. Results from a new aoristic approach highlight a weekly Monday peak in motor vehicle crime on one division of Nottinghamshire Constabulary. In another example a historically weighted temporal model helps to more accurately identify rising or falling crime rates.

4.1. INTRODUCTION

Chapter 3 (Data sources and software) identified some of the failings in police recorded crime data, and these become more apparent when analysis takes place. These shortcomings include the lack of temporal accuracy in the date and time of criminal incidents. This usually occurs because the actual time (and occasionally the date) of the incident is unknown to the victim. With the inherent inaccuracy of the source data, it is not surprising that this temporal aspect of crime analysis has received less attention in the academic world. Even within the police service, although active police officers often know *where* crime is concentrated, they may not necessarily know when it is happening with any accuracy, nor whether the localised crime rate is rising or falling. This chapter aims to rectify this, and the analysis presented here may allow them to monitor changing local temporal crime patterns obscured by other methods.

Temporal analysis of crime already takes place at national and regional levels. The annual publication of Home Office crime figures is a regular media event and local communities can compare the rise or fall in crime figures with the results for previous years at the county and national level. On a local level a crime analyst at a police division might plot the change in numbers of crime 'events' across the various beats of a station *if* the beat information is included with the crime record. It should be noted that the term 'event' has become accepted as a way of distinguishing the position of an individual observation within the study area (Gatrell *et al.*, 1996). Temporal analysis of crime patterns is valuable for a number of reasons. There may be value in detecting the change in crime level over time and moving the focus away from those areas with the highest concentrations in crime. Lower concentration areas may be interesting if the crime rate has suddenly risen or fallen, not in relation to other areas, but in relation to the historical pattern in that area. Simplistic spatially constrained non-temporal hotspot analysis may obscure the variation in areas with lower levels of crime. If changes are evident, then one explanation may be presented by Cohen and Felson's routine activity theory (Cohen and Felson, 1979) which argues that most crime requires the convergence in space and time of a motivated offender, a suitable target and the absence of a capable guardian.

Until the introduction of detailed spatially referenced crime data the police beat was generally the smallest resolution available for temporal and spatial analysis. This chapter aims to suggest a methodology for variable scale temporal analysis of crime data and in the process identifies some of the benefits and pitfalls which the user can expect from applying aoristic spatiotemporal process models. Examples from a police division of Nottinghamshire Constabulary provide real world case studies of the process in action. During the course of the study a significant temporal relationship was uncovered, unobserved using other analytical techniques.

4.2. TEMPORAL ANALYSIS

It has been recognised increasingly that data models should incorporate a time element to reflect the changing nature of the world (Raafat *et al.*, 1994). Much of the research considers varying methodologies of recording change within a temporal database, and the problems of how much historical information to retain and in what format (Langran, 1989; Langran, 1993; Peuquet and Niu, 1995). These papers contain some of the more thorough reviews of recent temporal analysis and temporal database storage.

Peuquet (1994) recognised that there are two possible approaches to temporal data storage; a feature-based (vector) approach, and a location-based (raster) approach. Vector models can be used to chart the movement over time of a specific geographical object, such as the advancing of a shoreline (Langran, 1993). Langran referred to this type of method as relying on 'amendment vectors' as can be seen in Figure 4-1. Raster models can be used to chart changes in more generalised data such as land use data derived from remotely sensed images (Lowell, 1994; Peuquet, 1994). This data model can be visualised as an incremental series of 'snapshot' images of the study area. Each image (which might, for example, be a remotely-sensed image) can be compared to previous and subsequent images to chart the change that has taken place between the image acquisition dates. Each cell within a separate image contains the value for the corresponding location at that time (Peuquet and Niu, 1995). Peuquet (1994) described the image as representing a 'world state' at a given time. However whilst simple to conceptualise and implement there are limitations to the 'snapshot' approach. Only the 'world state' is stored, and not the change that has taken place. The model is also incapable of informing the user of any changes that have taken place within the dates of the two snapshots, changes that did not persevere to the next data capture. As Langran (1989) observed; "snapshots store states, not changes."

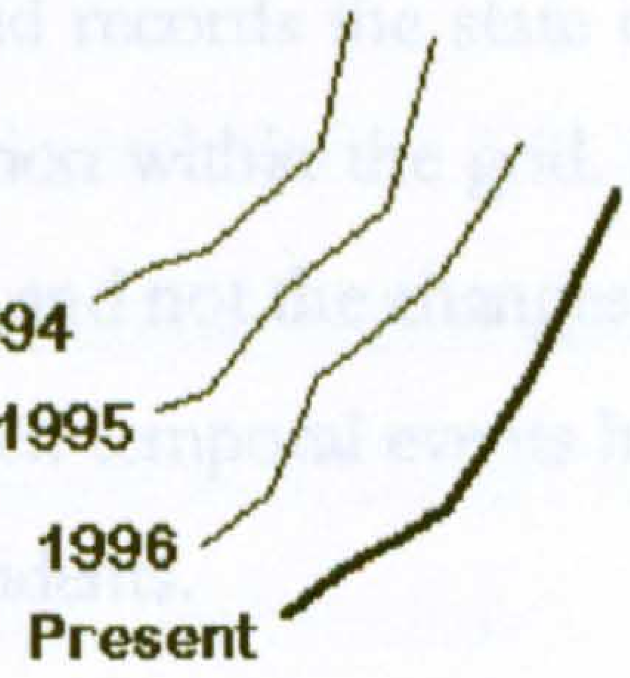


Figure 4-1 Langran's amendment vectors.

Source: Langran (1993).

In one article, Donna Peuquet presents a model for representing temporal change and likens the passage of time to an elastic line, with events appearing as knots at different places along the time-line (Peuquet, 1994). This is a useful analogy and allows us to conceptualise temporal events as having a start point, a duration and an end point. Temporal topological relationships can be defined with Boolean operators used to quantify the dynamics of any interaction between events (Peuquet and Niu, 1995). Different organisational approaches include time based methods where the temporal dimension is the main organisational factor, and the more straight-forward 'snapshot' approach presented here.

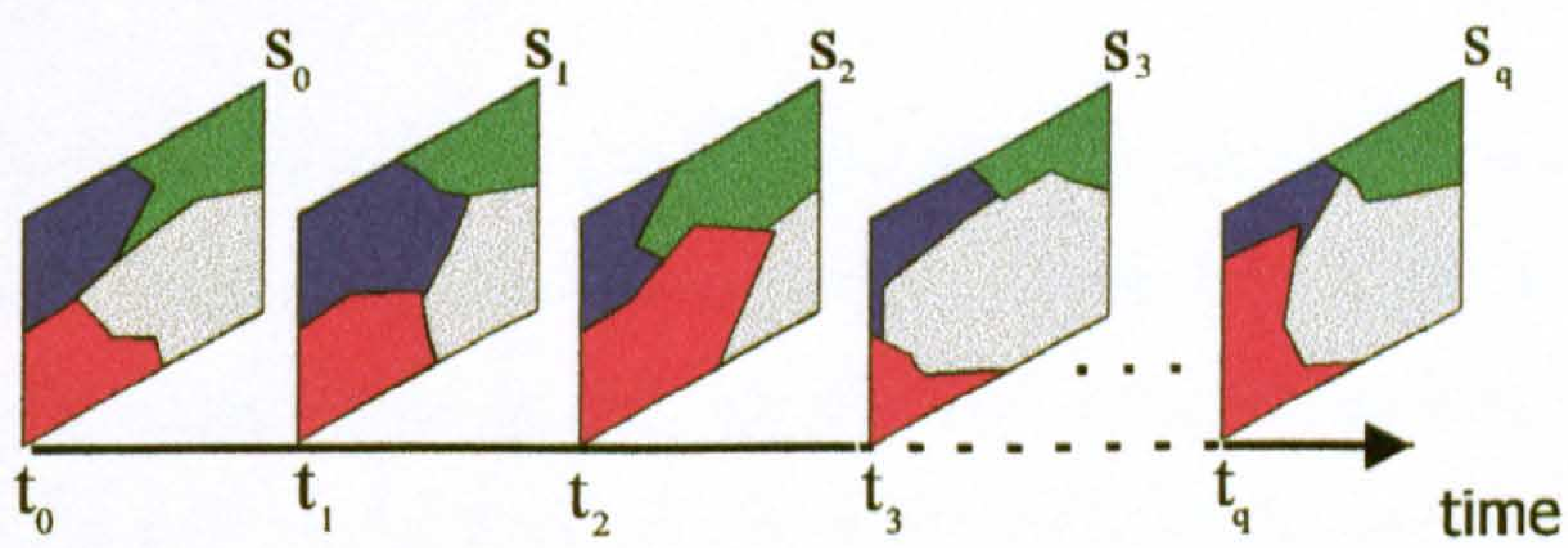


Figure 4-2 Snapshot temporal analysis approach.

The snapshot method is not only the most easily comprehended data model, it is also fairly easy to integrate into current GIS. A sequence of spatially-registered grids is employed to record the 'world state' at a given time. At different times the same grid is employed to record the 'world state' at another time and in this manner a collection of images is recorded, each at a unique time. Figure 4-2 shows S_i a snapshot of a given state taken at a time t_i. As Peuquet and Niu point out, the temporal distance between snapshots does not necessarily have to be uniform. Each pixel on the grid records the state of the desired variable at the chosen time, for just that location within the grid. It should be noted that each 'snapshot' records a given state and not the changes that have taken place. These methods are most effective when temporal events have a definite structure. This is rarely the case with crime incidents.

4.3. TEMPORAL ANALYSIS OF CRIME DATA

Temporal queries of crime databases are more involved because of the complex nature of crime recording. Unless the criminal is disturbed or captured, it is unlikely that the exact time of the offence will be known. Although rarely an issue for crimes like robbery (where victims tend to know when they were robbed) this is a problem for incidents like motor vehicle theft and burglary - crimes which are often identified as high prevention and detection priorities for police forces. Daytime burglaries can occur when the occupants are at work and are episodes which rarely take longer than a few minutes. When interviewed by the police the victims are unable to narrow the incident down to any time more specific than within a number of hours. Police crime recording practices reflect this by documenting a number of fields for the time of incident. Although field names vary from force to force, variations on *on_date, at_time, from_date, from_time, to_date* and *to_time* allow the crime record sheet to incorporate a range of possible incident times. This can vary from an exact event time to a number of weeks or longer, if a commercial premises has been burgled during a holiday period for example.

This creates problems when it is desired to search a crime database for events that occurred during a specific time period. A number of solutions to this problem are possible, however each has limitations on its functionality. During the course of this study three different methods of temporal search were considered; an averaging temporal search which averages the date and time fields, a rigid temporal search which only contains definite records within the search criteria, and an aoristic search which considers all records which *might* have occurred within the search criteria time. Aoristic is defined in the Shorter Oxford English Dictionary as:

> ***Aoristic**, without defined occurrence in time,* from*; **Aorist*** (SOED)*; one of the past tenses of the Greek verb, which denotes a simple past occurrence, with none of the limitations of the other [past] tenses.*

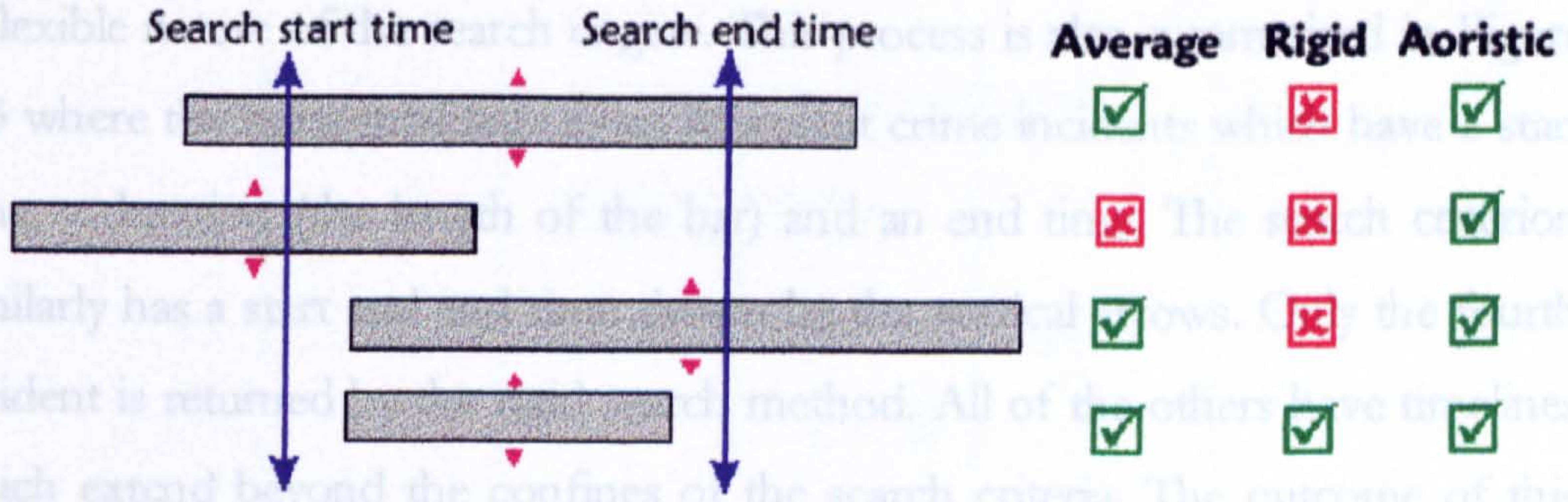


Figure 4-3 Summary of averaging, rigid and aoristic temporal search processes.

One of the simplest solutions is to take the average of the start and end event times. For example, if the victims of a burglary leave their house at noon and return to their burgled house at 6pm, the time of 3pm would be used by an averaging temporal method for any crime analysis. This process is summarised in Figure 4-3. Horizontal bars (labelled a-d) represent crime incidents which have a start time, a duration (the length of the bar) and an end time. The search criteria similarly has a start and end time shown by the vertical blue arrows. The small red markers in the middle of each event represent the location of the average along the time line. One of the limitations of the averaging methods is visible in the second incident from the top (b). Although a considerable amount of the second incident might have taken place within the search parameters, it is not included because the location of the average is just outside the search criteria. Averaging the date field has been used in studies of burglary repeat victimisation (Johnson *et al.*, 1997). Although computationally simple and an adequate solution for the longer time periods of repeat victimisation studies, this is a compromise answer and ties the event time to one possible time which, although the mean of the possible event times, is no likelier than any other time. A search based on fixed temporal windows adds emphasis to only one search period and denies the incident the chance to register with other equally applicable categories. One of the few advantages of this type of temporal search is that each event will be recorded at one definite time and the results of the temporal query will have the same number of incidents as the originating database.

Another possibility is to record those incidents which definitely fall within the search period. In this chapter this is called the rigid temporal search method. Search queries have to be constructed more carefully to reflect the more inflexible nature of the search engine. This process is also summarised in Figure 4-3 where the horizontal bars again represent crime incidents which have a start time, a duration (the length of the bar) and an end time. The search criterion similarly has a start and end time shown by the vertical arrows. Only the fourth incident is returned by the rigid search method. All of the others have timelines which extend beyond the confines of the search criteria. The outcome of this type of search method is a result with generally lower numbers than the originating database, but with a higher degree of accuracy in the temporal search.

Aoristic crime presents a particular problem in that it is temporally uncertain and can fall either side of the search boundary. The third approach, called here the **aoristic** search method, records crime incidents that *might* have occurred within the search time. Conceptually it is more complicated because a single incident can simultaneously register in a number of search categories if it covers an extended time period. In this manner an incident will appear in each time category where the slightest possibility exists that it might have occurred in that search block, however small a percentage of the incident time covered the search category. The resulting table will show a larger number of 'hits' than existed in the originating database, each 'hit' registering the *possibility* of an incident rather than a definite incident as in the previous method. This process is also summarised in Figure 4-3.

An extension of the aoristic search method is to address the problem of the larger number of 'hits' generated by the search procedure with crimes that have a time line extending a number of days, by proportionately reducing their influence. If a burglary happened at some point over a period of four days a **probabilistic aoristic** search would assign a value of 0.25 to each day to reflect the probability that the crime happened on that day. This probabilistic aoristic process is shown in Figure 4-4 where four different timelines are shown as horizontal bars (a – d). The amount that each can contribute to a daily search is shown in purple on the right axis, and the probabilistic aoristic totals for each of the four days is shown in red at the bottom.

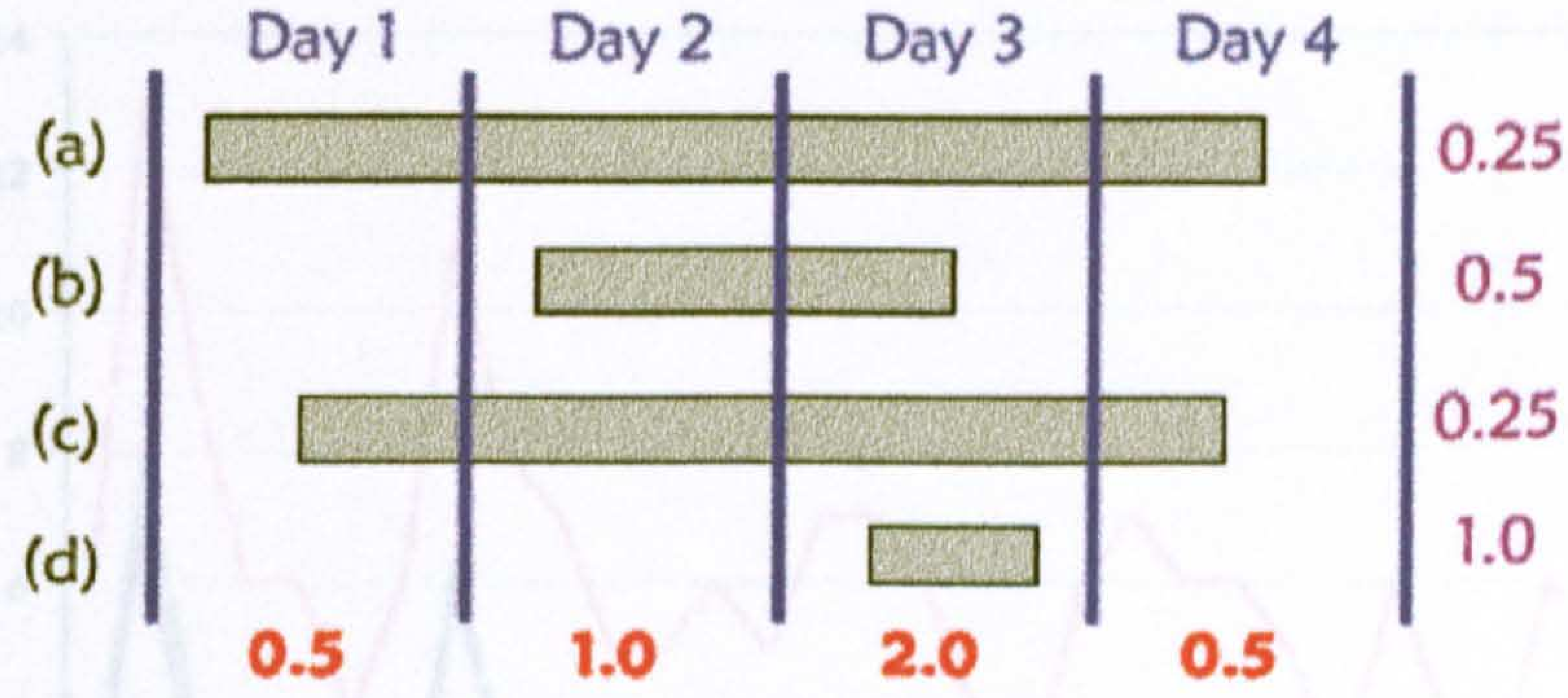


Figure 4-4 Probabilistic aoristic search method.

The diagram shows four crimes with variable length timelines (a-d). These crimes vary in time from (a) and (c) which stretch over a four day to two-day crimes (b) and single day incidents (d). The amount that each crime contributes per day to a probabilistic aoristic search is indicated by the purple figures on the right. For example crime (a) occurs over a four day period and therefore each day in a search would be allocated ¼ (0.25). Crime (d) definitely happens during day 3 and is assigned a value of 1.0. The red figures at the bottom of the diagram show the value allotted to each of the four days when the probabilistic aoristic method is applied.

The rigid and aoristic approaches offer more accurate and elegant solutions than the averaging method first described. The averaging temporal search method is acceptable for long term studies such as repeat victimisation and annual or monthly comparisons, but for shorter time scales it lacks accuracy and can easily generate erroneous results because it fixes the crime incidents at a totally arbitrary date and time. The rigid method presents the more precise means of identifying temporal events, while the aoristic method allows for the exploration of all possible events from the originating database, either as a straight aoristic search or by utilising the probabilistic aoristic process.

As their name suggests, the rigid and aoristic search methods offer different interpretations of the data in their results. The following factors should be considered in choosing which method to use. The size of the database generated might have repercussions for later statistical operations. The temporal vagueness of the original database is also important. The more accurate the temporal fields in the database, the less a method like aoristic search becomes necessary. Appendix C shows the full MapBasic code for generating Structured Query Language (SQL) aoristic and rigid searches of crime data.

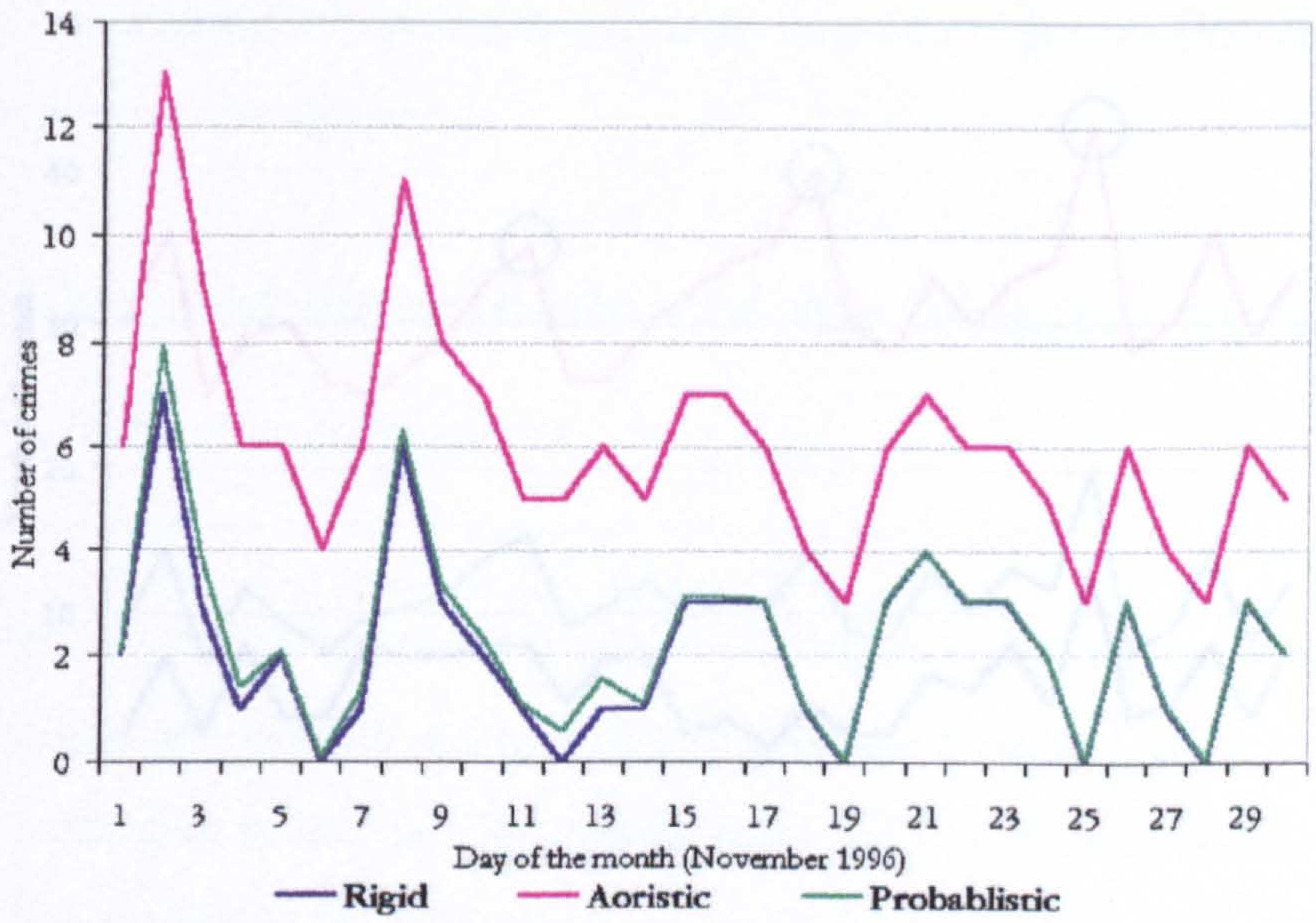


Figure 4-5 Rigid, aoristic and (aoristic) probabilistic assault search results.

The diagram shows the relationship between the different search procedures. Rigid and probabilistic share almost identical values except in the beginning of the month where three cases of domestic assault covering a period of years slightly adjust the probabilistic figure.

Figure 4-5 and Figure 4-6 were generated from a daily analysis of crime records from Trent division, Nottinghamshire Constabulary in November 1996. For a fuller description of Trent division see Chapter 3 (Data sources and software). Aoristic, probabilistic aoristic and rigid methods were employed to examine the daily crime rate for assault and motor vehicle crime. For definition purposes assault includes all counts of actual and grievous bodily harm. Car crime includes all counts of motor vehicle theft, theft from motor vehicle, and TWOCing (taking without consideration) - often referred to as joyriding.

Figure 4-5 shows the number of records recovered by the aoristic, probabilistic aoristic and rigid search methods when selecting all assaults. The curves show a strong positive correlation. The main difference between the results are the existence of three crime records which detailed cases of domestic assault lasting a period of years. In these circumstances the police often use one crime sheet to record an ongoing catalogue of incidents. Removal of these three incidents from the graph would bring the lines into almost exact correspondence. With this type of crime it is clear that the rigid search method is sufficient.

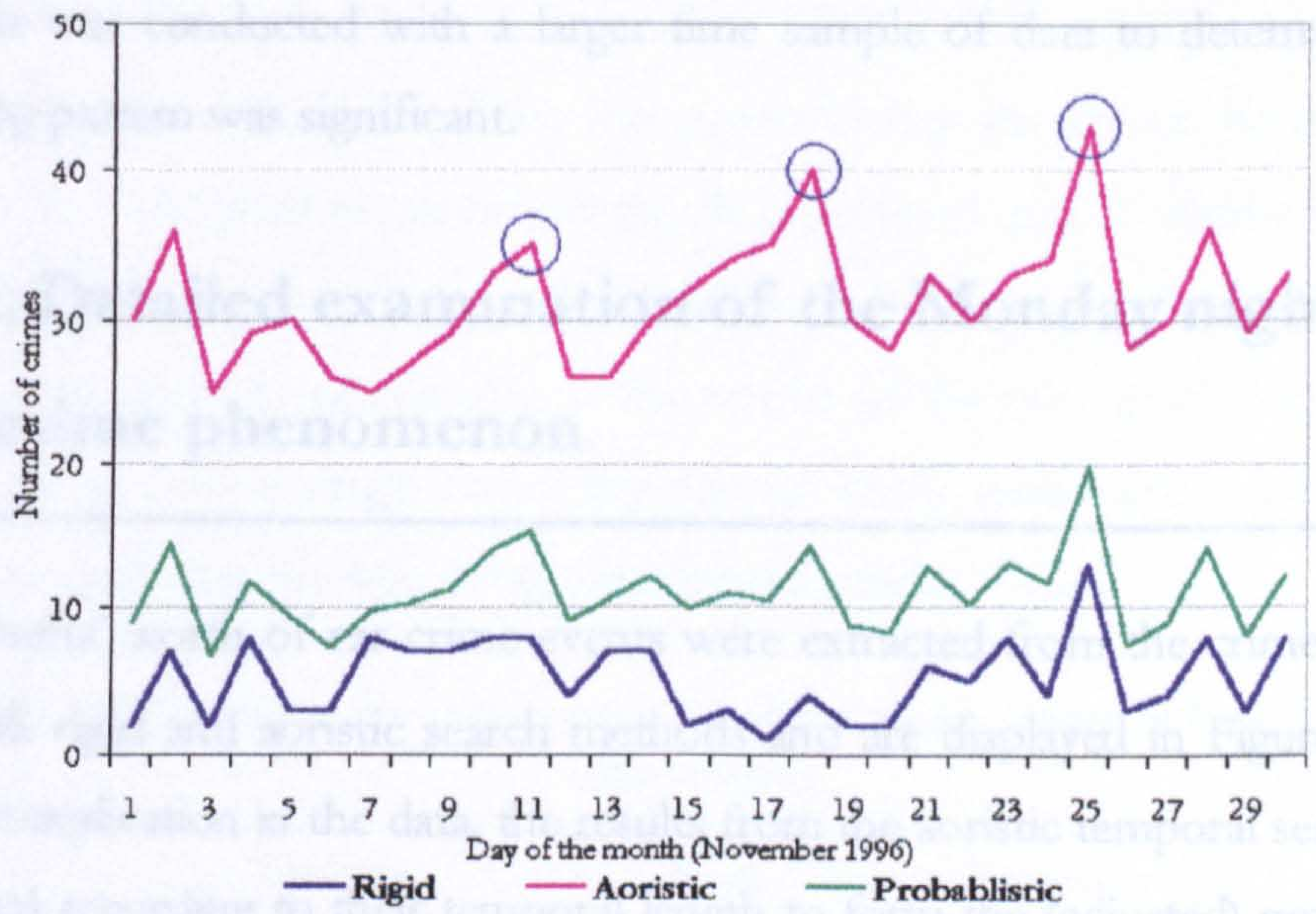


Figure 4-6 Rigid, aoristic and (aoristic) probabilistic car crime search results.

Note different vertical scale to Figure 4-5. Unlike the assaults diagram in Figure 4-5, the rigid and probabilistic graphs have significantly different values which reflects the greater number of uncertain vehicle crimes which might have occurred over a number of days. Circles indicate apparent peaks in the aoristic values on the 11th, 18th and 25th of the month.

Figure 4-6 shows the difference between rigid and aoristic search processes for car crime over the same study area and time. Note the difference in vertical scale from Figure 4-5. The absolute difference between the rigid and probabilistic values indicates the lack of temporal accuracy in police crime reports due to the inability to tie down the time of incident. Because these crimes may have occurred on one of a number of days the rigid method is unable to use them while the aoristic and probabilistic aoristic can include these incidents. This temporal inaccuracy is most evident with the apparent regular 7-day peaks in the aoristic data (circled in Figure 4-6). These occur on the 11th, 18th and 25th of November 1996 and are not detectable in the rigid data (with the exception of the 25th). The original data was re-examined in an attempt to discover why these Monday peaks existed. It was originally thought that the victims might have been away for the weekend and returned in the early hours of Monday morning to find their car stolen or items stolen from the vehicle. This was not the case and it appears from detailed crime record inspection that Sunday and Monday nights were simply the most popular with local car thieves in November 1996. This trend is not visible using the rigid temporal search method. A full statistical analysis was conducted with a larger time sample of data to determine if the Monday pattern was significant.

4.3.1. Detailed examination of the Monday night car crime phenomenon

Six months' worth of car crime events were extracted from the crime database for both rigid and aoristic search methods and are displayed in Figure 4-7. To prevent replication in the data, the results from the aoristic temporal search were weighted according to their temporal length to form the (adjusted) probabilistic aoristic set shown in Figure 4-7. In Figure 4-7 the blue line shows the full aoristic search result, and the black line identifies the (aoristic) probabilistically adjusted value.

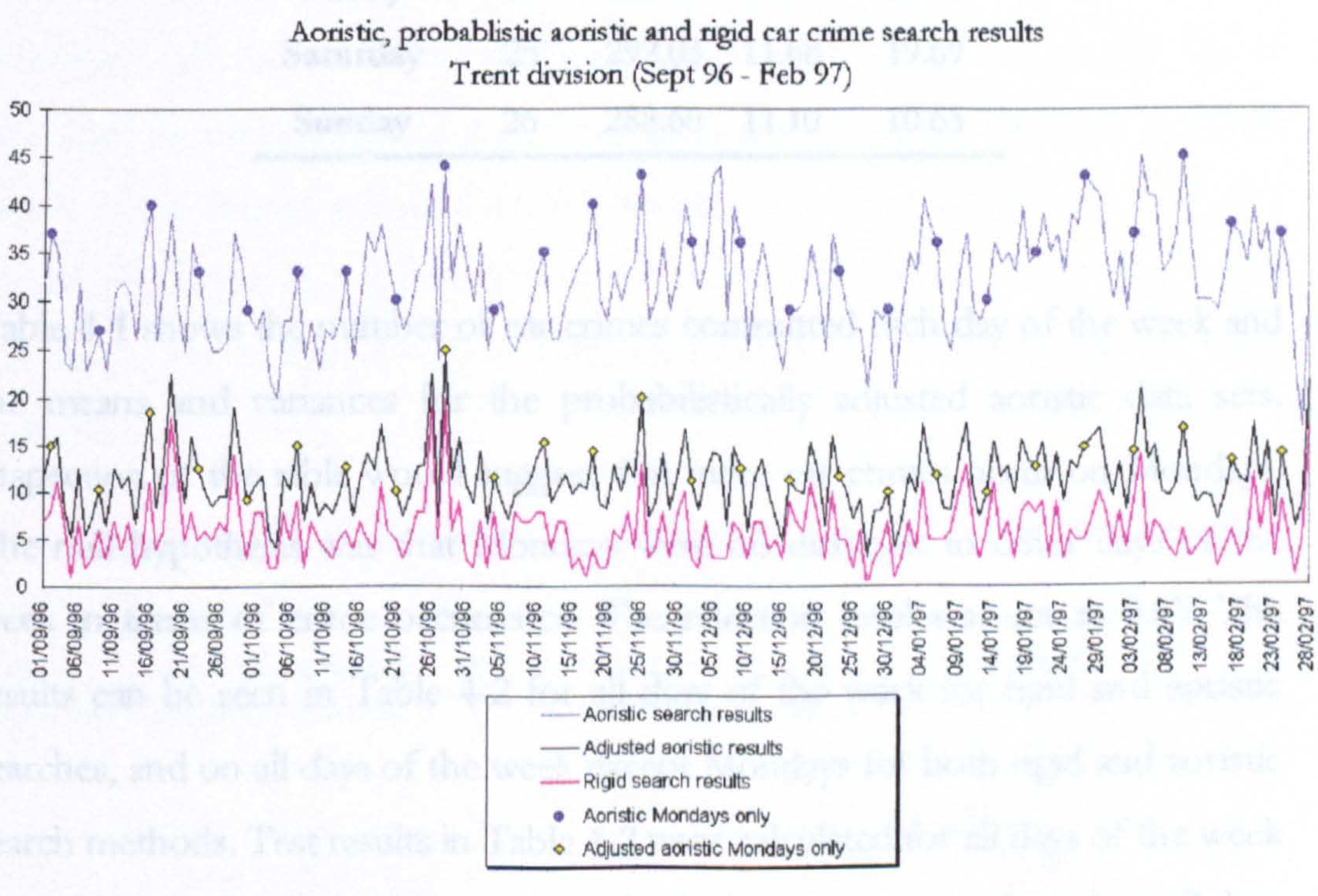


Figure 4-7 Daily count of car crime for Trent division for 6 month period.

The selection of appropriate statistical analysis was complicated by the unknown distribution of the data, the potential lack of sample independence in the aoristic data sets, and uncertainty as to the underlying data generation process. Two tests were used; a standard single factor exploratory Anova, considered by some authors (for example Alder and Roessler, 1962 p.245), robust under adverse distributional conditions; and the less powerful non-parametric Kruskal-Wallis test which is designed to use non-normally distributed data of diverse variability. The problem of sample independence remains, but in the author's view is not greatly infringed. A comparison of the results of the two approaches is given below. Four Anova single factor (by day of week) tests were calculated, two using the rigid data and two using the adjusted aoristic data.

Table 4-1 Basic counts by day of week using probablistic aoristic search of Trent division car crime records.

Groups	Count	Sum	Mean	Variance
Monday	26	355.87	13.69	12.06
Tuesday	26	268.35	10.32	14.03
Wednesday	26	272.80	10.49	6.83
Thursday	26	316.77	12.18	13.42
Friday	26	281.92	10.84	13.27
Saturday	25	292.03	11.68	19.69
Sunday	26	288.60	11.10	10.63

Table 4-1 shows the number of car crimes committed each day of the week and the means and variances for the probabilistically adjusted aoristic data sets. Inspection of the table would suggest that more car crimes occur on Mondays. The null hypothesis was that Mondays were no different to other days of the week in terms of crime occurrence. The rejection level was set at 0.05. The results can be seen in Table 4-2 for all days of the week for rigid and aoristic searches, and on all days of the week except Mondays for both rigid and aoristic search methods. Test results in Table 4-2 were calculated for all days of the week for rigid (1) and aoristic (3) searches, and a further test was conducted on all days of the week except Mondays - again for both rigid (2) and aoristic (4) search methods. The purpose of (1) and (3) was to test whether significant Monday peaks occurred in the data, and in (2) and (4) to show that the other days of the week were undifferentiated.

Table 4-2 Single factor Anova tests of search results from Trent division car crime events - September 1996 to February 1997.

Test number	Test of variance between groups	df	F value (calculated)	F_{crit} value (p = 0.05)	Probability level	Ho rejected
(1)	Rigid all days	6	1.94	2.15	0.08	No
(2)	Rigid not Mondays	5	1.45	2.27	0.21	No
(3)	Aoristic all days	6	2.80	2.15	0.01	YES
(4)	Aoristic not Mondays	5	1.02	2.27	0.41	No

There is no significance difference in either of the rigid search results (1) and (2). No variation in the pattern of crime, and specifically no Monday peak, has been detected by the rigid search method. Many records are excluded from the rigid search as they have timelines which extend beyond the daily search criteria (i.e. overnight). The aoristic method shows that Monday peaks are present (3). If Mondays are excluded (4), there is no significant difference between the daily means in the dataset and no pattern emerges. This study shows that there is a statistically strong Monday peak over the six month period. This result would not have come to light if only a purely rigid search for exact records had been performed.

Table 4-3 Single factor Kruskal-Wallis tests of search results from Trent division car crime events - September 1996 to February 1997.

Test number	Test of variance between groups	df	H value (Distributed as Chi Square)	Chi Square (p = 0.05)	Probability level	Ho rejected
(1)	Aoristic all days	6	17.69	12.59	<0.01	YES
(2)	Aoristic not Mondays	5	5.76	11.07	>0.40	No

The results of the equivalent Kruskal-Wallis test are shown in Table 4-3 for the aoristic data set alone. The same conclusions are reached with much the same significance levels as before. This would indicate some confidence in the results.

4.4. EFFECTS OF BOUNDARIES

Accurate temporal selection of crime incidents has to be delimited within a geographical framework and small scale boundaries such as enumeration districts, police sub-divisions or generated grids are adequate for the task. Grid patterns are easy to generate in programmable GIS. Software houses such as MapInfo even provide freeware packages for grid creation (currently only available for Lat/Long grid creation) and other functions which are available to download from their web site (http://www.mapinfo.com). Regular grids have an additional advantage in crime pattern analysis as they allow the generalisation of point data which preserves the anonymity of the original event locations. In this manner, results of crime analysis can be passed to outside agencies such as council crime prevention panels without compromising the sensitivity of the original data. For the purposes of this thesis a grid creation routine was written by the author in MapBasic. Whatever the chosen boundaries, it is not a requirement that they be the same shape or volume. A temporal analysis will compare crimes within one polygon with the same boundary at another time, and not with neighbouring locations. For example, if two grid squares are next to each other, one of them could be mostly parkland. A search of motor vehicle crime in a parkland square will usually exhibit a low score. This is not an issue unless environmental change leads to the building of new housing estates where the parkland used to be. Grids can however cause problems when the grid referencing system of the crime data is considered.

Within Nottinghamshire Constabulary most grid references are attached to crime records by querying the postcode and address and then attaching the AddressPoint 0.1 metre resolution Ordnance Survey grid reference from a central gazetteer. However the AddressPoint catalogue at Nottinghamshire Constabulary is not complete and any unknown locations have the Postcode Address File (PAF) 100 metre resolution grid reference merged with the crime event record as a back-up procedure. If a grid uses a similar 100 metre resolution, a number of crime events can find themselves on the **exact** location of grid intersections and either included in two (or more) grid results, or excluded from any grid location dependent on the GIS areal search procedure employed. This is due to the selection process used within the GIS. A standard spatial SQL term such as 'intersects' will find all points that are within a polygon or on the boundary of a polygon. If a polygon is the boundary between two regions, a similar process applied to the neighbouring region will also detect the same point. Using grid co-ordinates that are not rounded to the nearest 100 metres reduces this effect, although complete alleviation of the problem is only possible by reprogramming the areal selection tool in the GIS system used.

Figure 4-8 shows the study area, a one kilometre resolution grid square over part of West Bridgford, a suburb of Nottingham. A year's worth of data (1996) was used to produce a pattern of all crimes over 52 weeks. The grid corners were deliberately programmed to avoid the 100 metre resolution problem mentioned earlier. To examine the extent of this problem, the grid was converted to polylines and the database interrogated to find how many crime events intersected the lines of the grid. From a database of 8,267 records, 62 were found to exactly intersect the lines of the grid, only 0.8%. With a sufficiently large database, this is an acceptably low percentage, but caution must be exercised if much smaller grids with finer resolution are employed, increasing the number of lines and raising the probability of intersection.

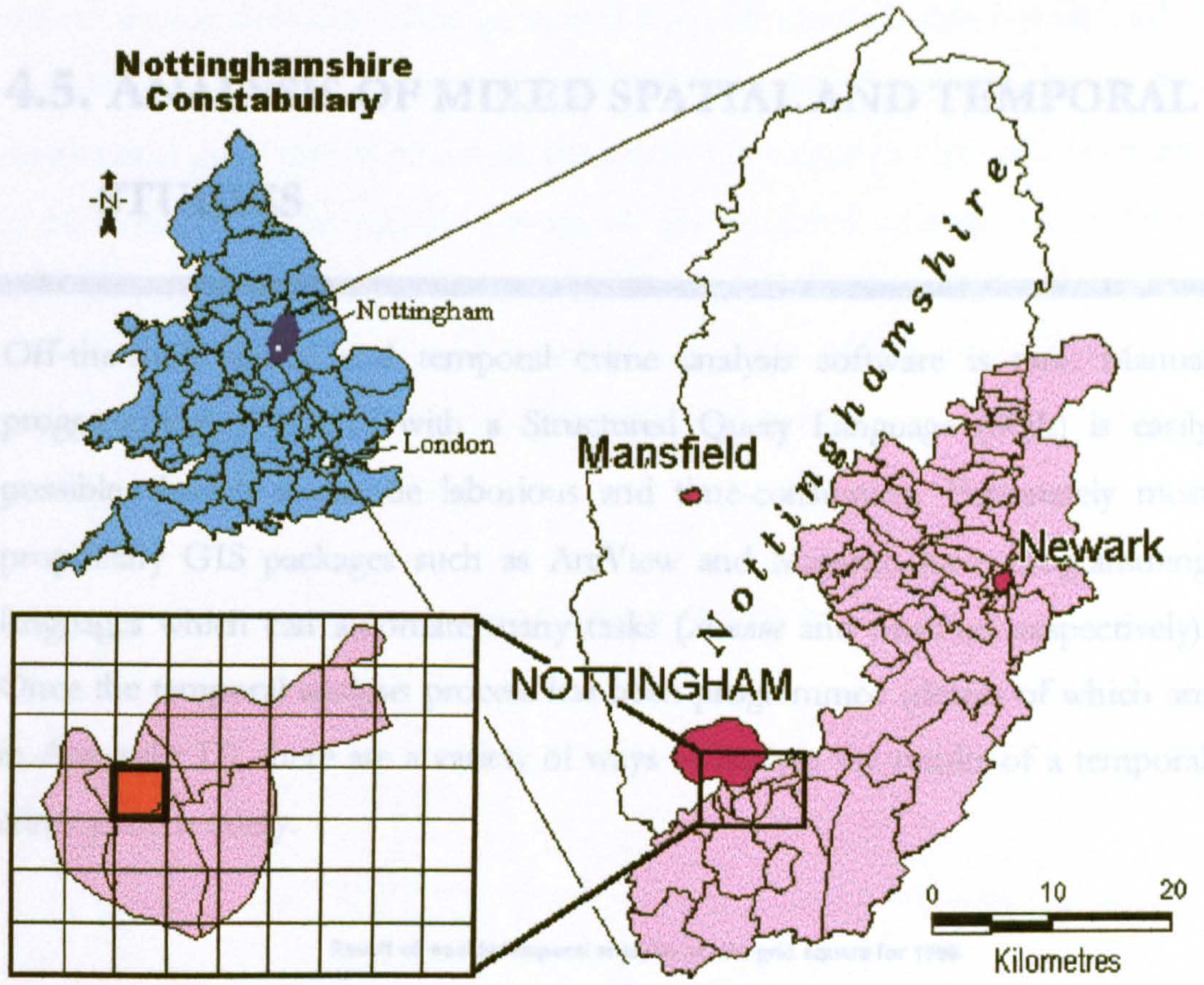


Figure 4-8 One kilometre resolution grid over West Bridgford station area, Trent Division, Nottinghamshire Constabulary, UK.

Study cell within the grid is outlined. The large purple map on the right shows the current Trent division beats - the smallest resolution for local spatial crime analysis prior to the use of GIS. The location of Nottinghamshire is shown in the blue map on the left.

4.5. ANALYSIS OF MIXED SPATIAL AND TEMPORAL STUDIES

Off-the-shelf spatial and temporal crime analysis software is rare. Manual programming of a GIS with a Structured Query Language (SQL) is easily possible, though it can be laborious and time-consuming. Fortunately most proprietary GIS packages such as ArcView and MapInfo have programming languages which can automate many tasks (*Avenue* and *MapBasic* respectively). Once the temporal analysis process has been programmed (details of which are in Appendix D), there are a variety of ways to analyse the results of a temporal crime pattern query.

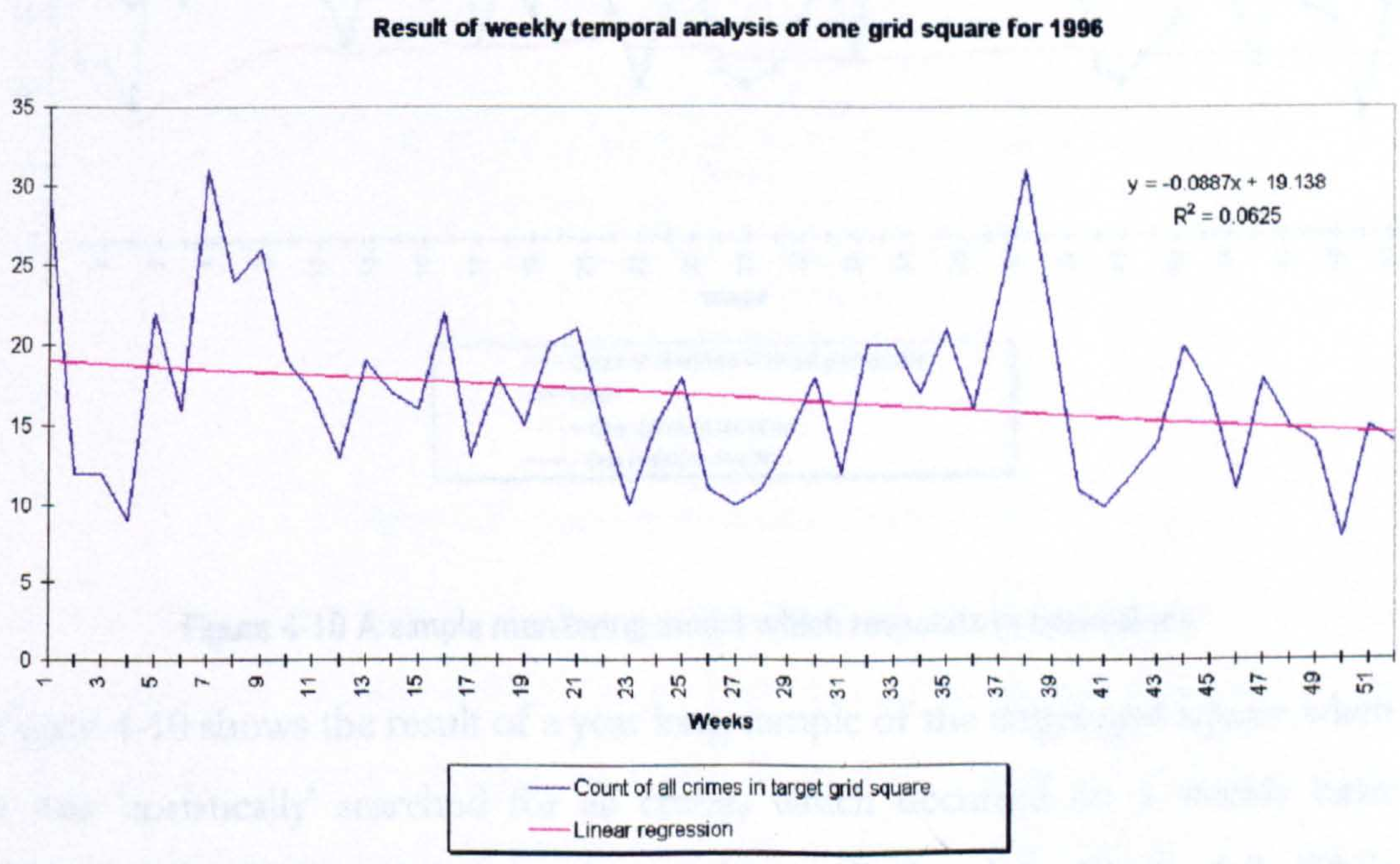


Figure 4-9 Simple regression of annual crime count of a single 1km grid cell.

Two realistic aims are to monitor the trend in crime rate for an area over time (often achieved using some form of regression analysis), and secondly to identify time periods of particularly high or low rates to enable further investigation of their causes. A simple example of this type of regression analysis is shown in Figure 4-9 where the aoristic count of all crimes in the cell identified in Figure

4-9 are shown with a trendline generated from Microsoft Excel. The data takes on added complexity because the grid or boundary framework has a two dimensional geographical structure, the amount of crime in each grid cell exists as the third dimension and the passage of time becomes a fourth dimension to display. During this study it was felt that the most effective means of analysing the data was by using the variety of techniques available within a database program. Each geographical unit can be examined individually and the change in crime rate plotted against time.

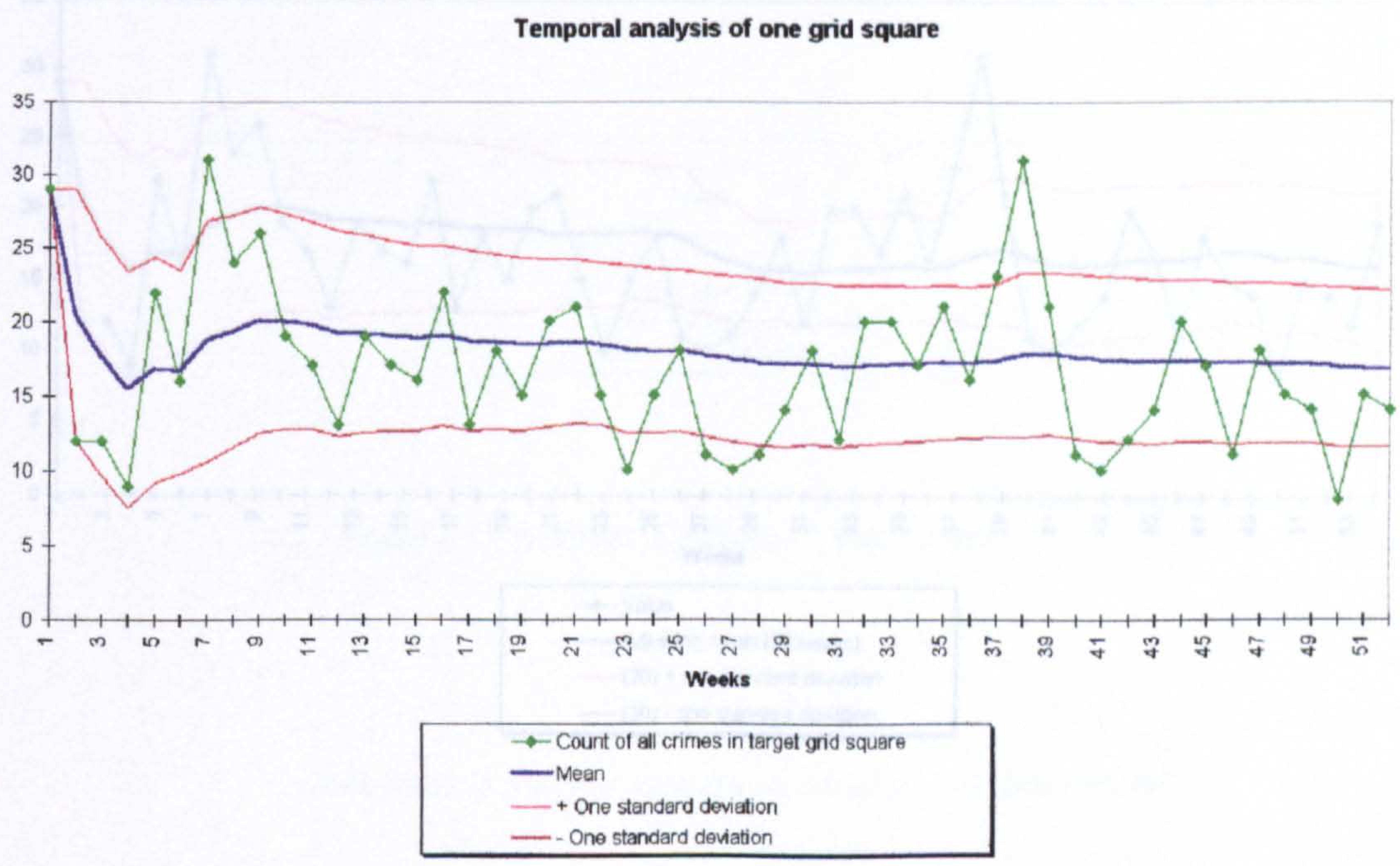


Figure 4-10 A simple monitoring model which responds to new values.

Figure 4-10 shows the result of a year long sample of the target grid square when it was 'aoristically' searched for all crimes which occurred on a weekly basis during 1996. This is the same data which is displayed in Figure 4-9. While regression analysis has applications in policework for mapping the general trend of the crime patterns in grid cells, one possible alternative could be an adaptive running mean. Figure 4-10 shows the count of all crimes in the target grid square for the same study period. It also shows the accumulating mean adjustment as each new value is received by the system. With every new value a revised standard deviation is calculated and two guidelines show the average plus and minus of one standard deviation. This adaptive process enables monitoring of crime rates to determine highs and lows on a continuously variable basis.

The advantage of this system is that a threshold can be set and an alert given if exceeded. These unusual values can then be examined to find the cause. A 'settling down' period can be seen at the start during which time the system is of limited value owing to insufficient values being included in the mean.

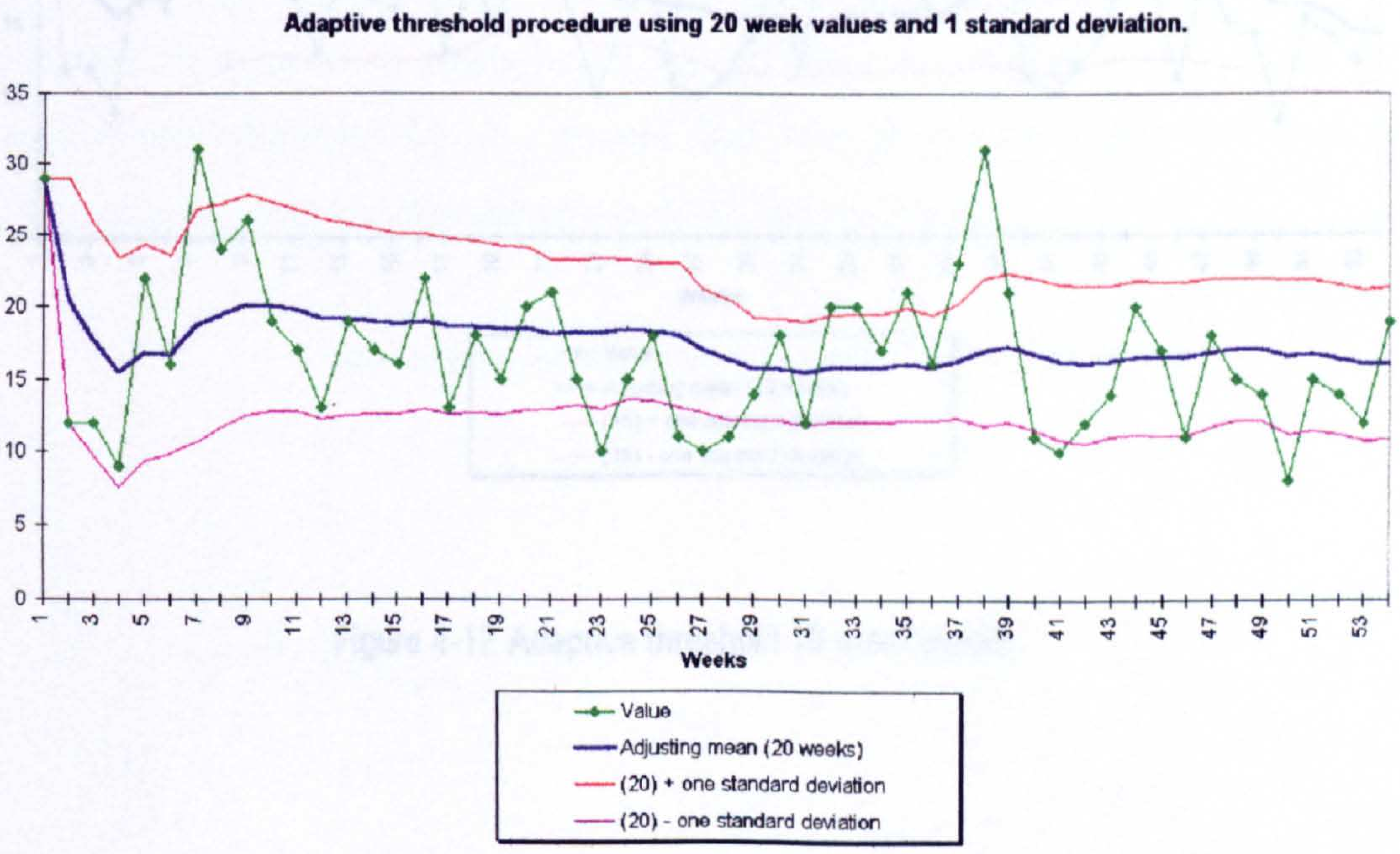


Figure 4-11 Adaptive threshold 20 week model.

An extension of this type of analysis can be the limiting of the number of weeks which influence the standard deviates. The process in Figure 4-10 allows the initial values to remain influential to the mean and standard deviates months after they were calculated. A time limit can be imposed to restrict the number of values entered into the calculation. This can be seen in Figure 4-11, Figure 4-12 and Figure 4-13 where the number of contributing values is restricted to 20 weeks, 15 weeks and 10 weeks respectively. The first ten weeks values are identical as each system includes the same values. After the 10th week has passed it can be seen that the ten week model reacts more rapidly to new values as they immediately contribute 10% to the mean and standard deviation calculations. The 20 week model is less reactive as each new value only contributes 5% to the equations. The 10 week model is more peaky as would be expected, while the 20 week model is smoother and slower to react to excessive values.

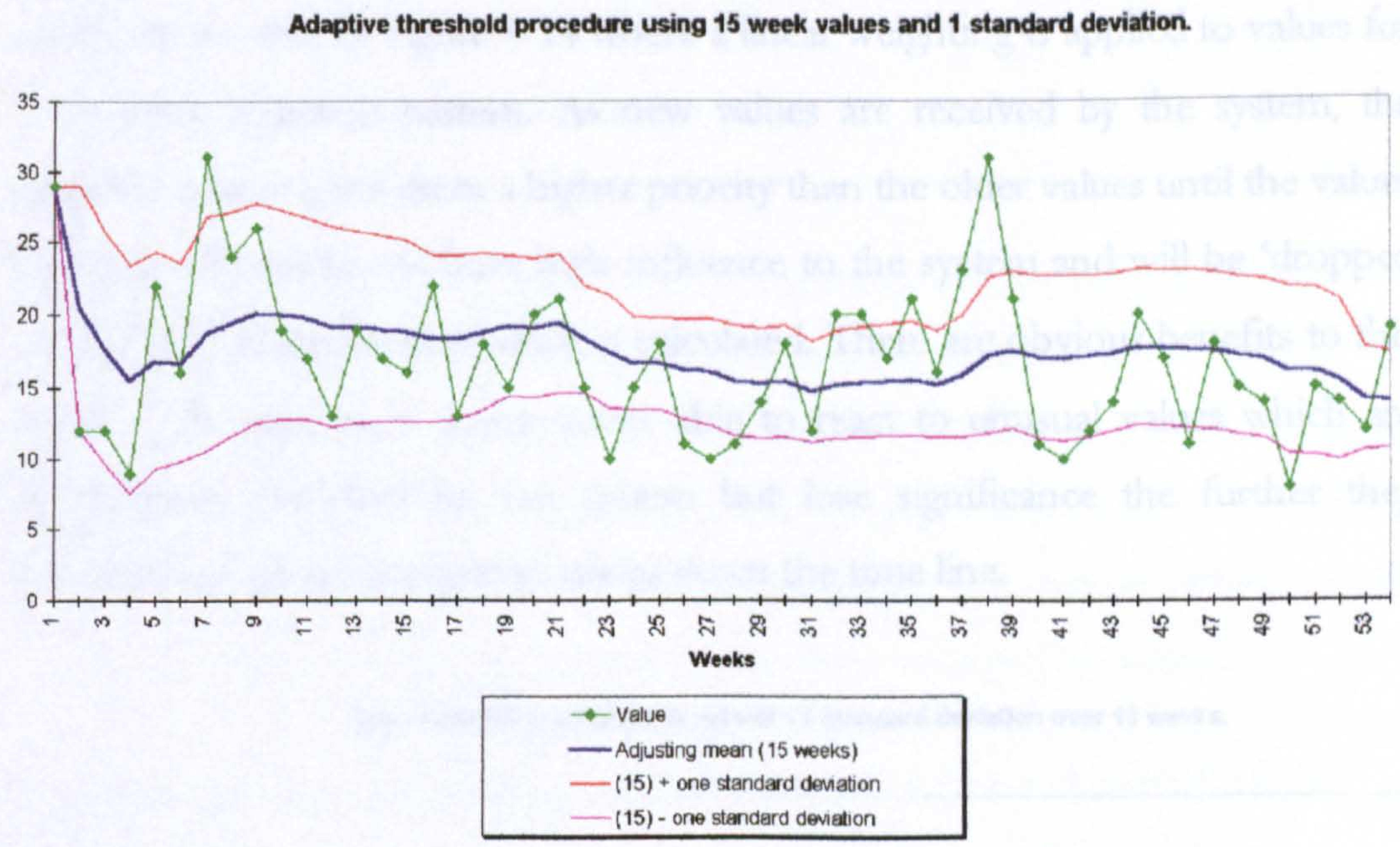


Figure 4-12 Adaptive threshold 15 week model.

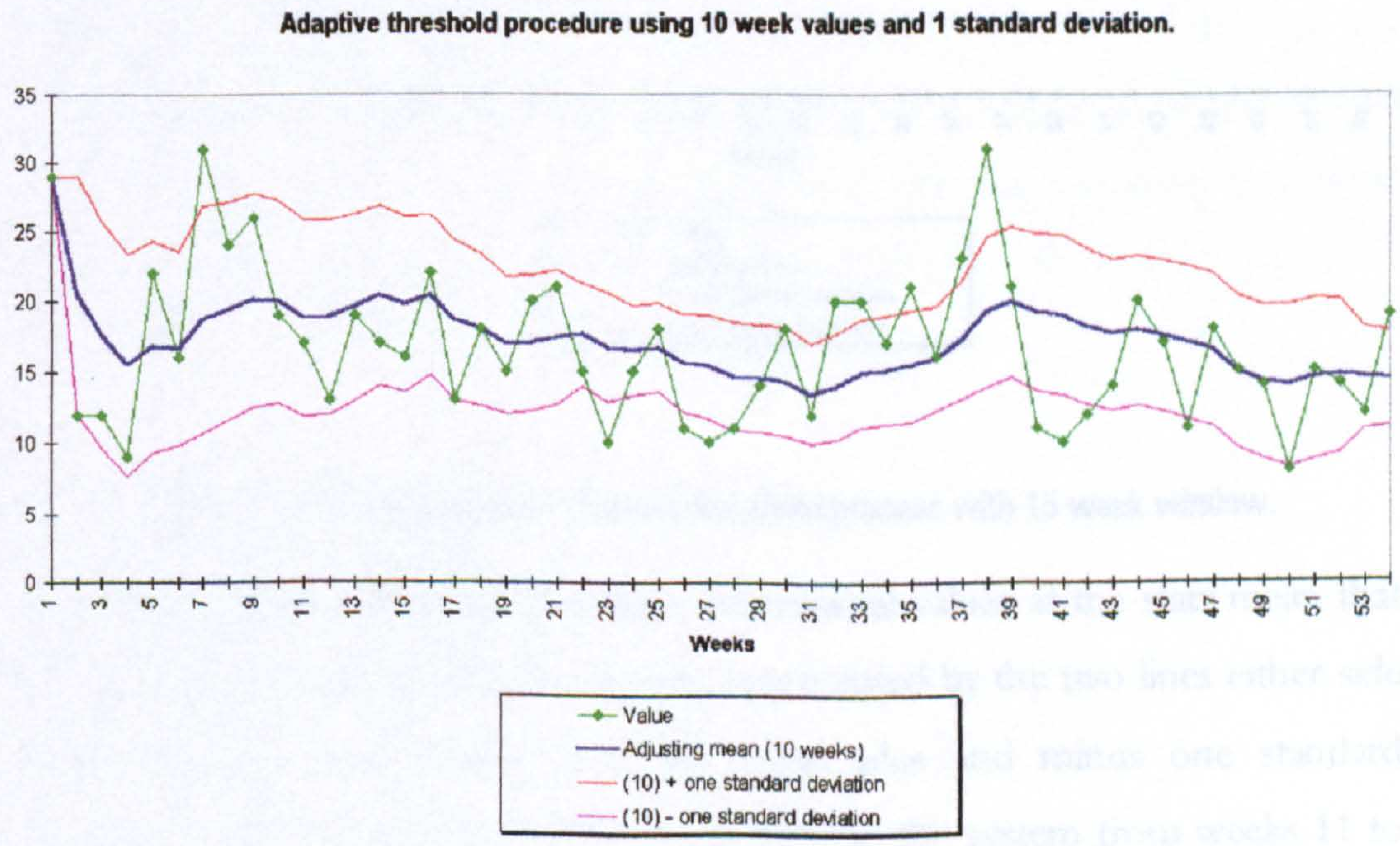


Figure 4-13 Adaptive threshold 10 week model.

A further extension of this process can be considered by using a linear weighting system over the most recent *n* values in the running mean, weighted in a descending linear pattern which gives emphasis to the most recent values. The result can be seen in Figure 4-14 where a linear weighting is applied to values for a 15 week historical pattern. As new values are received by the system, the adaptive system gives them a higher priority than the older values until the values which are 15 weeks old have little influence to the system and will be 'dropped off the end' when the next value is calculated. There are obvious benefits to this system. The process is much better able to react to unusual values which are immediately absorbed by the system but lose significance the further they proceed away from the newest values down the time line.

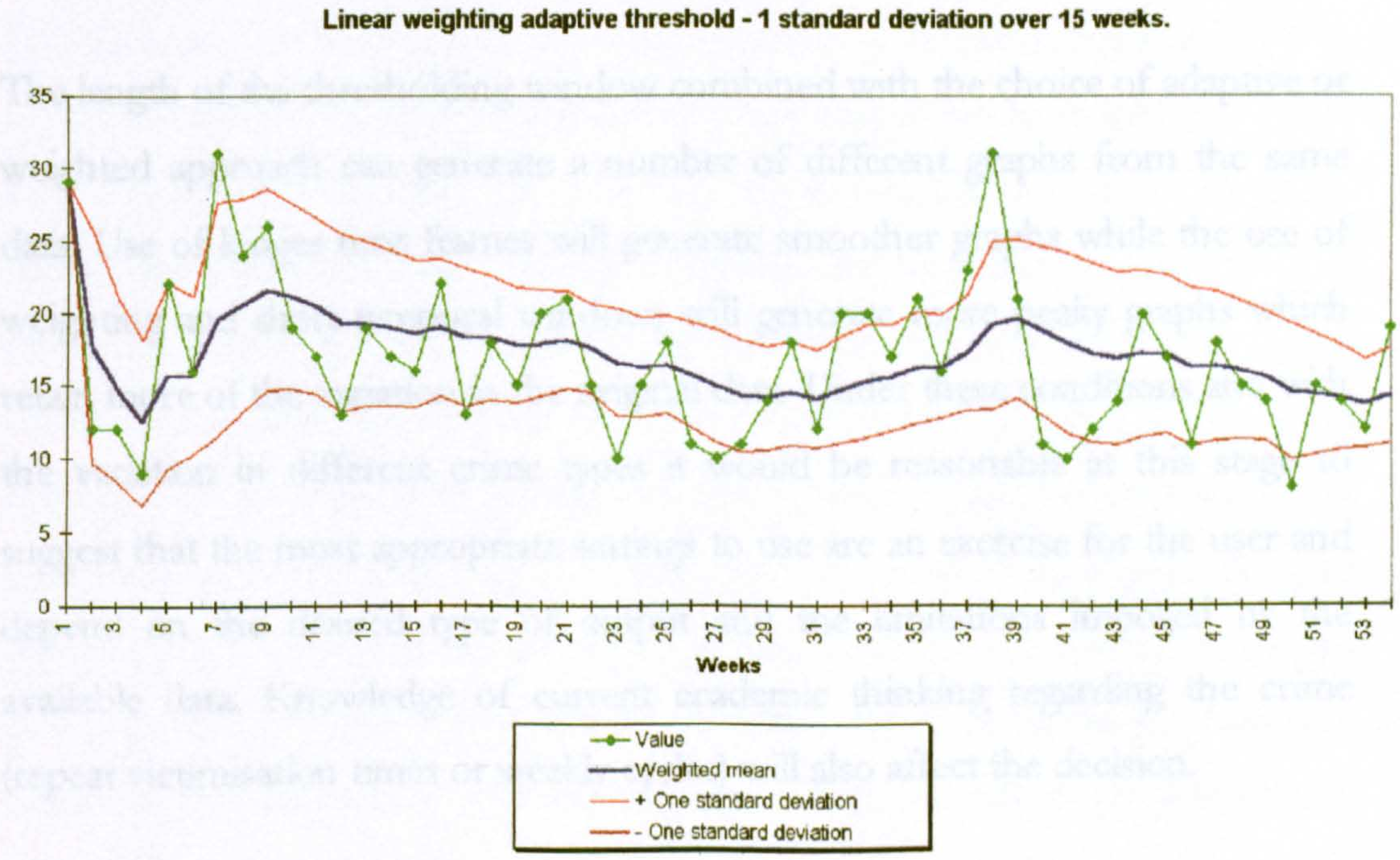


Figure 4-14 Linear weighting adaptive threshold process with 15 week window.

After the initial 'settling-down' period, the unusual values at the start mean that the model has a high standard deviation, represented by the two lines either side of the central mean which show the mean plus and minus one standard deviation. Once more stable values are entered to the system from weeks 11 to 35, the standard deviation rapidly reduces and the lines converge on the mean, indicating that the unusual values are exerting less influence. Weeks 37 to 39 are again unusually high and the system adapts rapidly to reflect this, though the standard deviation is dropping by the end of the graph as more stable values are received by the system. Although a single standard deviation is used here as the threshold, any measure of weighted sample variability could be employed.

4.5.1. Applying the models

The question remains as to the choice of a suitable threshold. This chapter has shown a weekly peak in motor vehicle crime and a police practitioner might use this weekly frequency as the basis for an adaptive thresholding process. Chapter 6 (Identifying repeat victimisation) will show a longer time period for burglary repeat victimisation extending to a number of weeks, and various different types of crime might necessitate a different choice of thresholding approach.

The length of the thresholding window combined with the choice of adaptive or weighted approach can generate a number of different graphs from the same data. Use of longer time frames will generate smoother graphs while the use of weighting and short temporal windows will generate more peaky graphs which retain more of the variation in the original data. Under these conditions and with the variation in different crime types it would be reasonable at this stage to suggest that the most appropriate settings to use are an exercise for the user and depend on the desired type of output and the limitations imposed by the available data. Knowledge of current academic thinking regarding the crime (repeat victimisation times or weekly cycles) will also affect the decision.

4.6. DISPLAY OF RESULTS

4.6.1. Computer animation

The results of the analysis of one grid cell were presented in Figure 4-14, however if a one kilometre resolution search area is used for the whole of West Bridgford it creates a 9 by 7 grid of 63 squares. This creates a problem for presenting the results. Animation has been suggested as the solution to many GIS problems involving time as a variable (Dorling and Openshaw, 1992; MacEachren, 1994). This is a field where animation might be particularly useful. GIS packages can be programmed easily to output selected screens as image files of screen captures. A repeated number of these images can be moulded together into a FLI or FLC file, an AVI file or a number of other computer movie/animation formats. Various programs are available to perform this operation, including a number of free or shareware products like Dave's Targa Animator (available through HENSA – see internet bibliography). It is beyond the scope of this chapter to describe the full animation process, but a suggested animation sequence might be the definition of a map surface with the grid hidden over the top of it. As the sequence (and therefore the weeks) progresses grid cells only become 'active' if they are either above or below the tolerance threshold. Effective use of colour could code cells as one colour for being above and a different colour for being below the threshold (for a more thorough treatment of the complete animation process see McCullagh, 1998). This technique has a number of advantages. It is an effective means of displaying a large quantity of numerical information in an easily understandable format. It is also an efficient means of explaining the temporal aspect, and utilises a grid base which preserves the anonymity of the original data. Modern processors and algorithms have speeded up the animation process to the extent that a several 100 frame animation can be created on a domestic PC in a few hours. Most of this time is required to generate the individual images. Only a few minutes are needed to run the images together to form an animation.

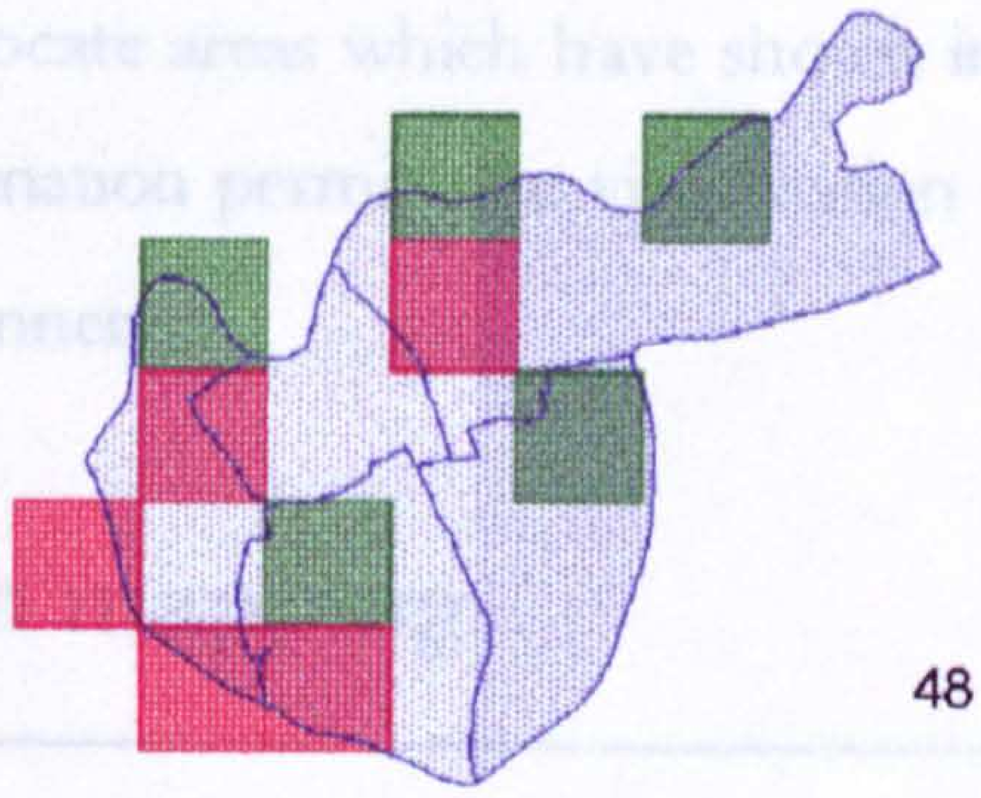


Figure 4-15 Single frame from aoristic_crime_analysis.avi animation.

This process was completed for all grid squares in the study area and an animation constructed in Adobe Premiere version 4.2. Figure 4-15 shows a single image from the animation. The number in the bottom right of the image indicates the week being displayed (in this case 48). The blue hatched area shows the five beats of West Bridgford sub-division. Red and green squares indicate the one kilometre square grid cells which are displaying an aoristic crime count for week 48 which deviates more than one standard deviation from the mean. If the cell is shaded green then the new value is more than one standard deviation below the mean, and if the new value is more than one standard deviation greater than the mean the cell is shaded red.

The full animation is available on the CD-ROM which accompanied this thesis. The filename is AORISTIC CRIME ANALYSIS.AVI and the details and path of the file are described in the first pages of this thesis. This file should be recognised by a Windows 95 operating system or newer. If this is not however possible, the file is also available in FLC format. This format is playable with AAWIN.exe which is also included in the CD-ROM. Details at the beginning of the thesis. The animation filename for the FLC format is AORISTIC CRIME ANALYSIS.FLC.

The use of animation permits the user to see a snapshot of the variation in crime patterns across the geographical area of West Bridgford on a week-to-week basis. This method of visualisation allows the two dimensional spatial dimensions to be combined with the changes in crime pattern (third dimension) over time (fourth dimension) in a unique way which would otherwise be extremely difficult to see clearly using any other technique. The ability to watch the crime pattern unfold permits the user to identify areas with deteriorating crime problems, or to locate areas which have shown improvements in criminal activity. This use of animation permits the visualisation of four dimensional data in a comprehensible manner.

4.6.2. Choropleth mapping

Without the ability to generate animation, the temporal change can still be summarised in a more limited but accessible means by showing the number of threshold crossings within each grid square on a simple choropleth map. Figure 4-16 shows a number of choropleth maps of West Bridgford station beats, a sub-division of Trent division. Figure 4-16a shows the summary of all crimes on the sub-division for 1996. This type of map is one of the standard range of simple distribution maps occasionally employed in police work (usually at the beat level and not gridded as here), which can mislead the user as to changes in more complex crime patterns. Figure 4-16b and Figure 4-16c depict the number of times over the year that a grid square's aoristic temporal count exceeded the weighted running mean by more than one standard deviation above (Figure 4-16c) and below (Figure 4-16b). The combination of these two images in Figure 4-16d shows areas with complex rapidly changing weekly crime rates, and also the areas with more stable crime event patterns. The temporal analysis in Figure 4-16 allows a much more detailed interpretation of the temporal crime pattern and shows that changes in pattern are occurring in very different locations from those of major crime. The picture is much more complex than that given by the single display of total crime in Figure 4-16a.

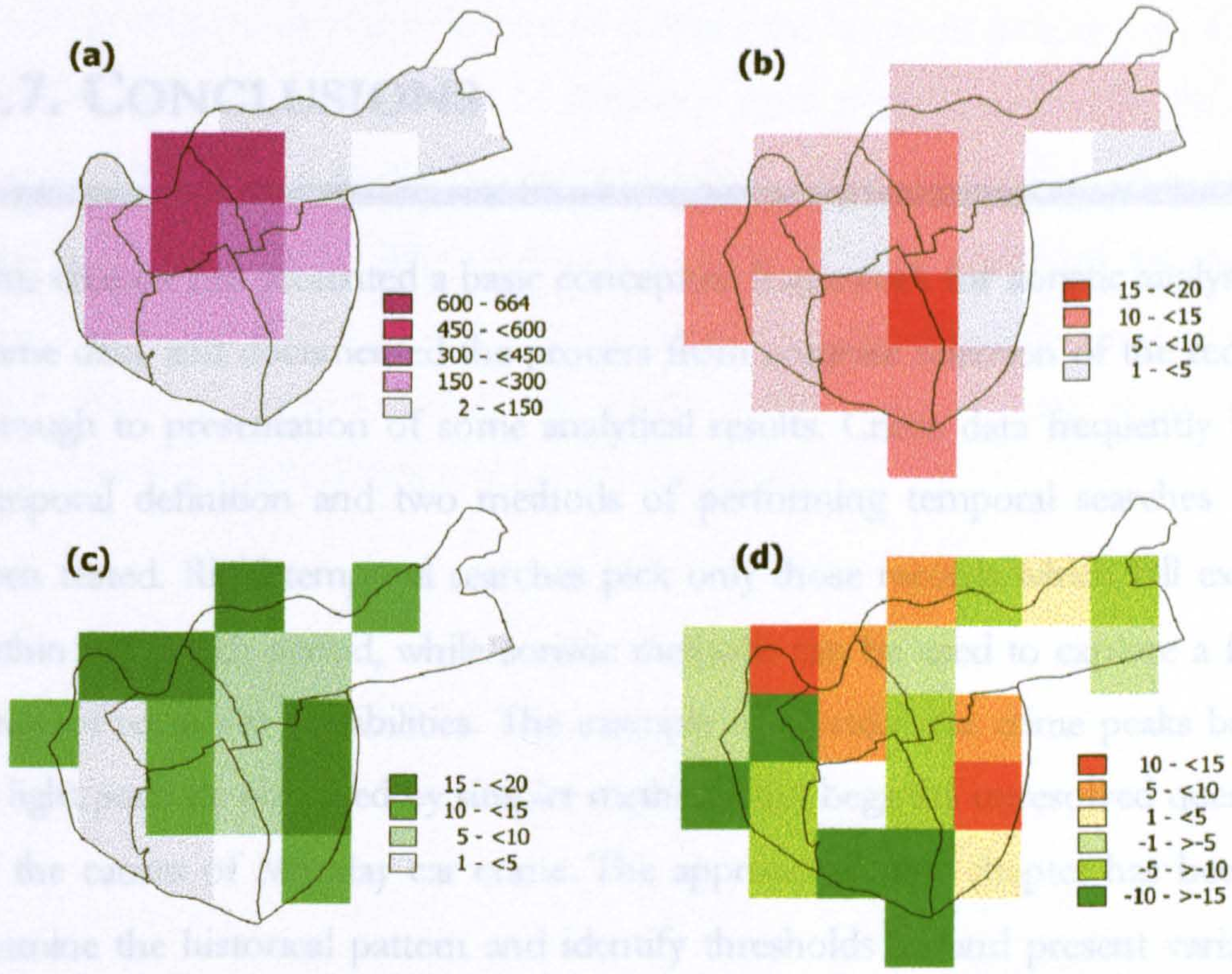


Figure 4-16 West Bridgford station beats overlaid with one kilometre square grids showing general crime and temporal analysis summaries.

(a) shows a standard mapping practice using totals for the year aggregated within one kilometre square grids. The aoristic search generated data set used to determine weighted running mean threshold results is shown in (b). Here the number of times the upper *threshold (above one standard deviation) has been exceeded during the year is mapped.(c) shows the same data as for (b) but a count of the number of times the* lower *threshold (below one standard deviation) has been crossed. In (d), the difference between (b) and (c) to show the predominance of upper or lower threshold crossings. A negative class on the map indicates a larger number of significant reductions in crime (gradations of green) and a positive class indicates significant increases (gradations of red).*

4.7. CONCLUSIONS

This chapter has presented a basic conceptual framework for aoristic analysis of crime data, and documented the process from accurate selection of the records through to presentation of some analytical results. Crime data frequently lacks temporal definition and two methods of performing temporal searches have been tested. Rigid temporal searches pick only those records which fall exactly within the search period, while aoristic methods can be used to explore a fuller range of temporal possibilities. The example of Monday car crime peaks brings to light patterns obscured by simpler methods, but begs the unresolved question of the causes of Monday car crime. The approach in this chapter has been to examine the historical pattern and identify thresholds around present variation outside which abnormalities can be evaluated.

The use of an aoristic search technique can be enhanced by the application of a probability function which allocates a value to each day in the search depending on the number of days across which the temporally unspecific crime spreads. This allows the final total of aoristic values to equal the total number of all the crimes but reallocates the values in a more accurate manner to the alternative processes of averaging or rigid searches. This process retains all of the advantages of the other systems (total is equal to the actual number of crimes) and avoids many of their disadvantages of temporal inaccuracy or inappropriate averaging.

The technique is limited in some ways. The probabilistic aoristic approach is limited to the minimum temporal unit being applied. If the search process is using a unit base of days, then a crime which happened sometime between 2358 hours and 2358 hours on the next day would register in both days though only two minutes of the crime occurred in the first day. This is a limitation of the process irrespective of the time unit applied and must be considered by the user. Decreasing the size of the temporal unit (hours instead of days) reduces but does not alleviate the problem.

Aoristic crime analysis is a method of charting the historical pattern over variable small areal units, independent of influence from neighbouring cells. Each cell can be plotted to examine the rise and fall of crime compared with a variable threshold over time. Temporal trends *in extrema* can be examined for patterns and can be correlated against other intelligence gathered at the police station. The adjusted aoristic search procedure, added to adaptive thresholding techniques, provides a new tool for crime analysis.

5. Spatiotemporal differences between crime and incidents

This chapter extends the temporal analysis of crime developed in the previous chapter by using seasonal decomposition methods and autocorrelation functions to elicit greater understanding of the temporal pattern of crime. This picture of crime is complicated but enhanced by the addition of incident data which shows a markedly different temporal and spatial distribution to the crime data. The chapter ends with a first attempt at a predictive system by building on the temporal information extracted from the earlier analyses.

5.1. INTRODUCTION

This chapter develops the temporal analysis employed in the previous chapter and enhances the analysis with the inclusion of an incident database generated from calls to the police and recorded by the Command and Control system throughout the Nottinghamshire Constabulary area. This type of data is commonly referred to in Australia and the United States as 'calls for service'. This data has been made available for analysis in the final year and includes every incident for just over a one year period from January 1997 to March 1998. A more complete description of the incident data is included with Chapter 3 (Data sources and software).

The incident data and crime records available for this research corresponds temporally for three months only (January 1997 to March 1997). This overlap still has great value as few researchers have access to both data sources, and only one article published in British academia could be found which utilised both sets of data. Alex Hirschfield and Kate Bowers used incident data as an indicator of social cohesion in an area (Hirschfield and Bowers, 1997) in their examination of crime and disadvantaged areas in Liverpool.

This chapter aims to examine what, if any, relationship exists between recorded crime and incident data with regard to crimes of disorder and related incidents. Events of a violent or disorderly nature appear in both the incident and crime databases and a relationship may exist between them. There would appear to be an intuitive link between these types of incident and crime which is more difficult to envisage with other types of crime. For example, it might be difficult to expect any type of incident to be a precursor of motor vehicle theft. Both disorder incidents and assault criminal investigations take up the time of the police but may display different temporal and spatial patterns. An objective of this chapter is to describe these variations. The chapter will also go on to examine if it is possible to employ incident data as a predictor of crime in the short term. One reasonable hypothesis is that incidents that happen at one time of the day are precursors of crimes that occur later in the day. It would be in the interests of the police if these reactive crimes could be predicted and prevented.

It should be borne in mind that this type of analysis has a number of significant limitations. The data sets are limited in temporal extent, covering only a period of 90 days. The two data sets are also not independent, but due to the nature of both data sources, it was impossible to identify accurately which incidents generated which crime reports. The full incident data contains an essential cross-reference to the crime database, but this critical field was not made available to the research. Some incidents generate crimes, some incidents do not result in crimes and some crimes are discovered directly by the police and never appear in the incident logs. With this complication in the data set accurate separation of the crimes and incidents was impossible.

This chapter employs a number of different temporal analysis techniques applied to a study area in Mansfield, North of Nottingham, where both crime and incident data are available. This chapter uses incident data extracted by Phillip Lewton of the School of Geography, University of Nottingham for his MSc. dissertation (Lewton, 1998). Although both Phil Lewton's study and this one share basic similarities in choice of subject matter and area, this chapter advances the study of the crime and incident data interaction by approaching the work with different temporal and spatial tools and building on the areas highlighted by Phil Lewton as worthy of further investigation. The chapter concludes by examining the results of the analysis, and identifying possible avenues of research in the future.

5.2. TIME SERIES AND SEASONALITY

Time series data is a set of recordings of a variable taken at a set interval and may contain intensity fluctuations over perhaps weekly or monthly periods. This fluctuation in the data set is termed seasonality. The analysis of time series is a complex process:

> *If we ignore the question of seasonality, we may make the error of assuming that a series is not seasonal, when in fact it is. On the other hand, if we automatically adjust for seasonality without first analyzing the series to see if it is seasonal or not, we may make the error of adjusting for nonexistent seasonality. (Block, 1984 p.3)*

By ignoring the first possibility of a seasonal influence in the data, we run the risk of making decisions based on data which might have a dominant seasonal pattern where knowledge of that pattern might affect the decision-making process. Alternatively error may be introduced to the analysis by assuming seasonality to a series that is not. This can lead to inaccurate forecasting of future trends. This can cause the model to 'overadjust' for seasonality and control for a seasonal variability which does not exist.

To analyse a time series correctly and avoid these types of error it is necessary to consider (1) if the series contains a seasonal component, and (2) what, if any, is the effect of the seasonality. There are three approaches to answering the question; 'Is this series seasonal?', and they share a number of similarities. ARIMA (AutoRegressive Integrated Moving Average) and component seasonality are the two main methods, and Fourier analysis is a third, less-used technique in this type of analysis. The two main approaches model the seasonal variability but the first approach incorporates the seasonal variability into the final analysis and the second emphasises a separate description of the seasonal fluctuation and the remainder of the data set.

The first approach incorporates the seasonal fluctuation into a complete model, and one of the most common methods is termed ARIMA. The second approach emphasises the breaking of the series into its seasonal and corrected 'components' and is referred to as the Seasonal Component method or the Census X-11 adjustment. The latter name originates from a seasonal adjustment program first introduced by the US Bureau of the Census. It has since become a widely-used seasonal adjustment standard. These two methods do show similarities in their mathematical approach and each can do much of the work of the other. For example, an ARIMA model can contain seasonal and non-seasonal terms that could be thought of as components. The differences between the ARIMA and component methods exist more in the means of application and the way they are interpreted (Block, 1984).

ARIMA

The ARIMA approach emphasises the forecasting of future patterns by describing and modelling past events. MacCleary and Hay (1980) emphasise that descriptions of each component of a time series will not necessarily produce a good forecast, even if each separate description of the components may be good. The ARIMA method does not separate the individual components, but aims to describe the stochastic processes of the entire series. The assumption that one observation follows another with a certain frequency is the basis of this stochastic model, though the application of this in a crime data environment is questionable. Measures that can affect the crime rate include: socio-economic conditions (Allen, 1996), architectural style (Coleman, 1985; Newman, 1972), unemployment (Elliott and Ellingworth, 1996), policing style (Bottomley and Coleman, 1976; Morgan and Newburn, 1997; Reiner, 1997), social cohesion (Hirschfield and Bowers, 1997), and even temperature (Field, 1992). Few of these variables can be realistically controlled in a crime series analysis. This author believes that the ARIMA method is more suited to situations where a time series is influenced by a smaller number of independent variables. The method is also suited to longer time series of years (Chatfield, 1989) and perhaps not the 3 months of data available in the Mansfield study.

COMPONENT SEASONALITY

The component definition of seasonality can be described as 'any cyclical or periodic fluctuation in a time series that recurs or repeats itself at the same phase of the cycle or period' (MacCleary and Hay, 1980). This approach aims to measure and remove the seasonal component from the rest of the time series. The technique is based on the theory that a time series has three components:

1. A trend/cycle component consisting of the long-term trend and any non-seasonal but regular fluctuations,
2. The seasonal component as described above,
3. The irregular component, which consists of everything else, including error.

A seasonally adjusted series is a series that retains the long-term trend and the residual variation (irregular component) but has had the seasonal factor removed. There are two main methods of calculating the seasonal factor. A multiplicative seasonal adjustment produces a seasonal component which is a factor by which the seasonally adjusted series is multiplied to reproduce the original series. Essentially the method estimates seasonal components that are proportional to the overall level of the series. Observations without seasonal variation have a seasonal component of 1. With an additive seasonal adjustment, seasonal adjustments are added to the seasonally adjusted series to obtain the observed values. This process attempts to remove the seasonal component from a time series so that other characteristics of interest in the data set that may be 'masked' by the seasonal component can be examined. Observations without seasonal variation have a seasonal component of 0.

The three components of the seasonality method (trend, seasonal component, and seasonally adjusted values) can be seen in Figure 5-7, Figure 5-8 and Figure 5-9 beginning on page 116, applied to recorded violent crime in Mansfield.

FOURIER ANALYSIS AND SPECTRAL ANALYSIS

A third method for identifying repeat cycles in data is Fourier analysis. Less popular in Geography, Fourier analysis has traditionally been the domain of electrical engineers and geologists (Davis, 1986), physicists and meteorologists (Chatfield, 1989). Spectral analysis is a modified form of Fourier analysis which makes it suitable for stochastic processes similar to the problems in this chapter, whereby a time series is decomposed into a sum of periodic functions and an error term. Spectral analysis is considered to be different from other time-series techniques (Chatfield, 1989). Instead of analysing the variation from one time point to the next, it analyses the variation of the series as a whole into periodic components of different frequencies. Spectral analysis attempts to approximate a function by a sum of sine and cosine terms to produce a periodogram. The periodogram is an unsmoothed plot of spectral amplitude (usually plotted on a logarithmic scale) against the period (wavelength or frequency). A plot of power versus harmonic number is also referred to as a power spectrum. A power spectrum is a good indication of cyclic occurrence. Low frequency plots can indicate a smooth series comprised of 'white noise' for some data, while peaks indicate the 'harmonics' of the series. A better indication of the true periodic structure of the time series can be seen when the periodogram is smoothed with a moving average. A number of processes are available for this (termed 'filters').

5.2.1. Autocorrelation and cross-correlation temporal analysis

As well as the search for a seasonal component in time series data, it might also be informative to compare a sequence with itself at successive positions and locate the maximum correspondence between the two identical sequences. One could also measure the degree of similarity between the sections of sequential data. This is the function performed by two different aspects of time series analysis (TSA). The two main techniques used are autocorrelation which compares the entire sequence with itself at all possible positions, and cross-correlation which compares two different sets of time sequences (Davis, 1986). For both types of analysis, the degree of similarity or dissimilarity is determined at all positions (known as match positions) on a scale of –1.0 (maximum dissimilarity) to 1.0 (maximum similarity). An example of this temporal analysis technique is shown with a simple pattern of values in Figure 5-1.

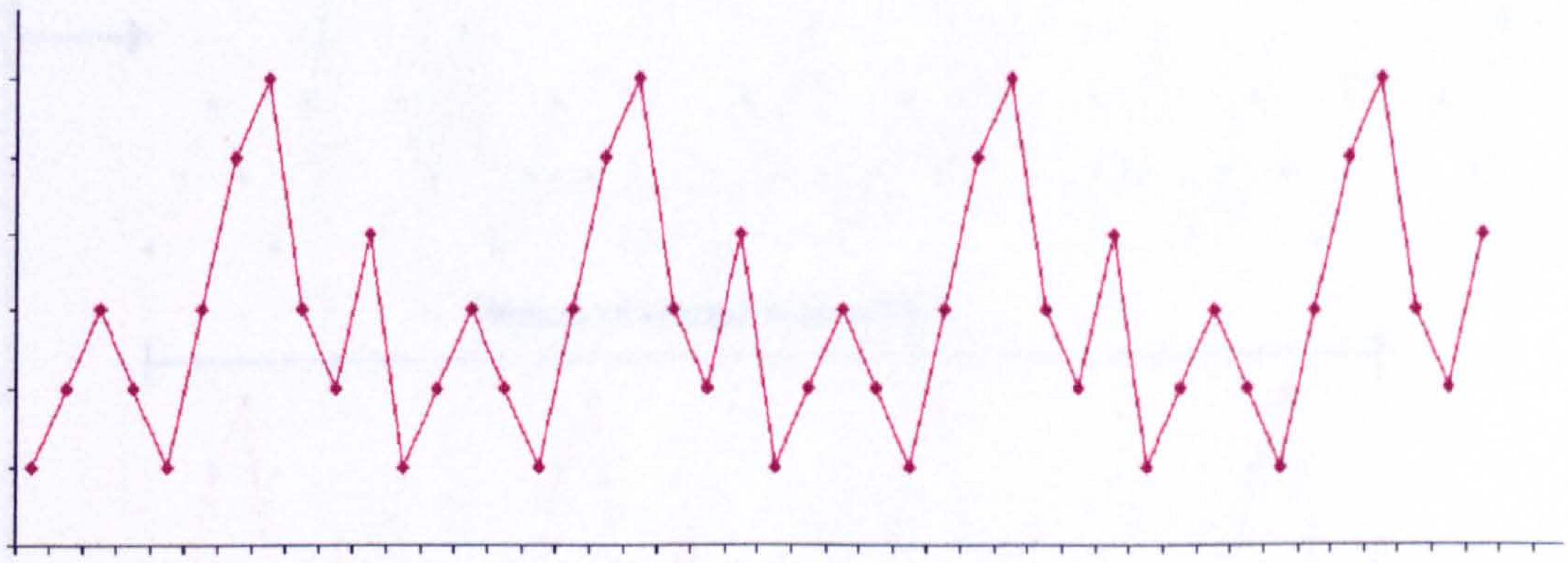


Figure 5-1 A sequence of nominal values that repeat after 11 readings.

Figure 5-1 shows a temporal sequence that appears to show a repetitive pattern of values along the x-axis that represents time. It is important to note with temporal analysis of this type that the values must be recorded with the same time interval. This is to say that the time gap between values on the x-axis must be the same. Autocorrelation works by sliding a copy of the sequence against itself and comparing the values at different lag times. This is shown in Figure 5-2 where the top sequence is an exact copy of the target sequence. The autocorrelation function lags the top sequence in a succession of values. Figure 5-2 shows the two sequences at match position 5 (the match position is also referred to as a lag position).

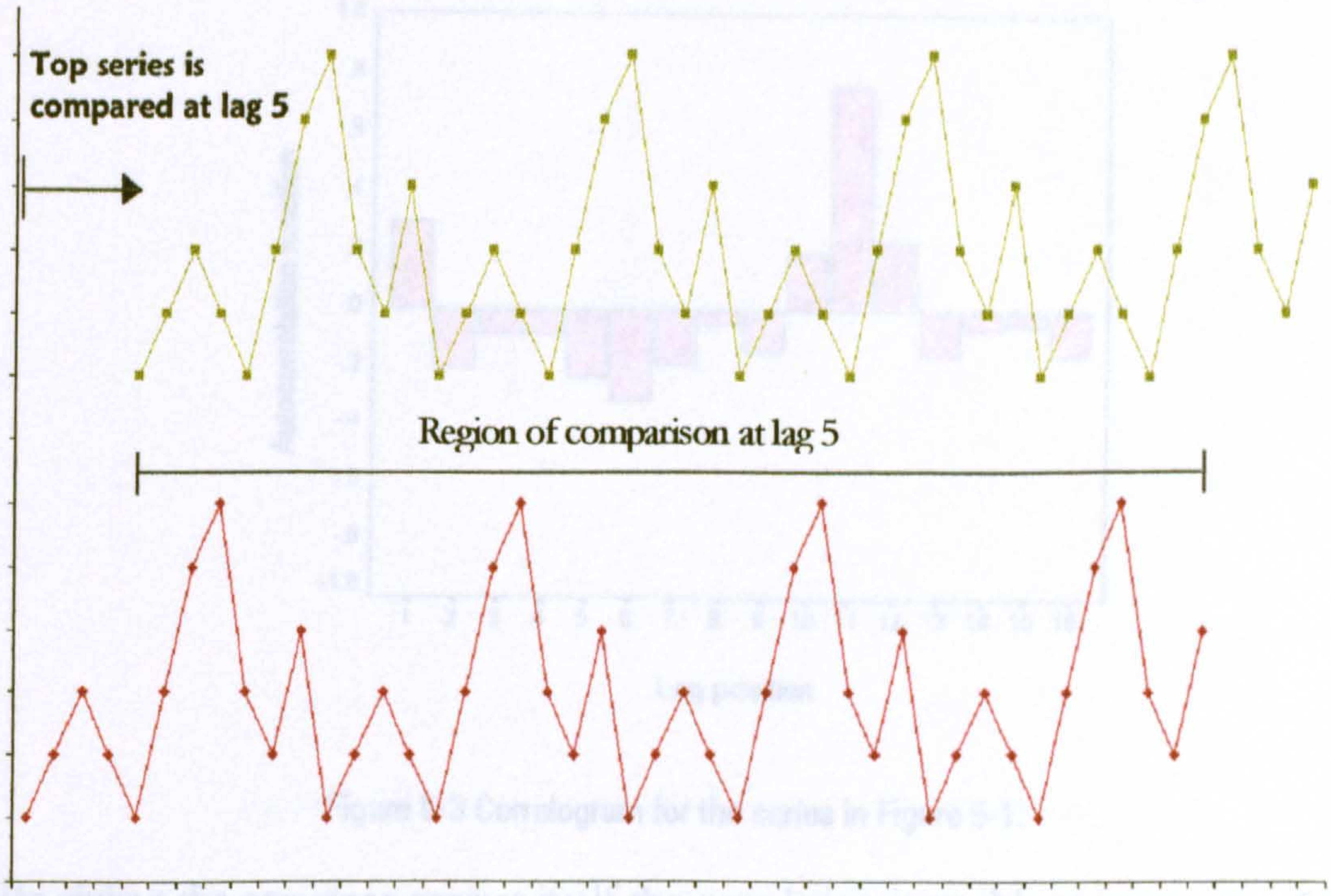


Figure 5-2 The same sequence from Figure 5-1 compared against itself at lag (match) position 5.

Once the autocorrelation has been calculated at every position it is possible to plot the results on a correlogram, which is a diagram of the correlation plotted against the lag positions. This correlogram can reveal characteristics of the time series (Davis, 1986). Figure 5-3 shows the correlogram for the series shown in Figure 5-1. As we would expect with a repetitive sequence the value at lag 11 shows a significantly high correlation. The sequence in our example time series repeats itself after 11 points.

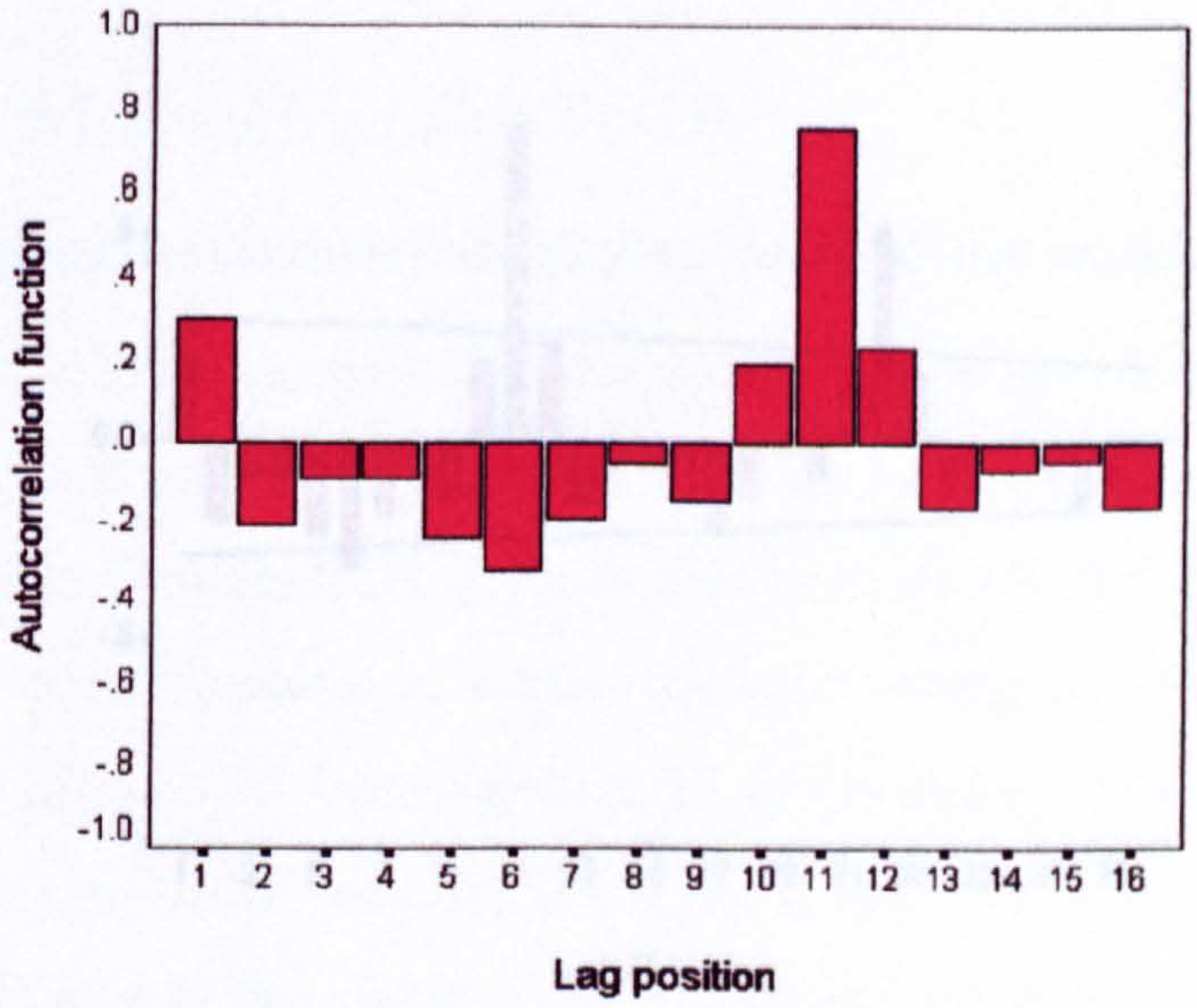


Figure 5-3 Correlogram for the series in Figure 5-1.

By sliding the sequence against itself the correlation is unable to compare a full spread of values. This is shown in Figure 5-2 where the region of comparison shows the first four of the bottom sequence and the last four of the top sequence are not considered completely. A feature of the autocorrelation function used in SPSS to generate Figure 5-3 is the inclusion of a weighting factor to account for the lack of overlap with series at longer lag times. As the lags increase, the number of overlapping values which can be analysed reduces. The linear weighting compensates for this. The reduction in the number of contributing values, and therefore the lack of a complete matching pattern explains why the correlogram is unable to show an autocorrelation function of exactly 1.0 for the matching sequence (Lag 11 in Figure 5-3). This is evident when the autocorrelation function is calculated for a longer time sequence (Figure 5-4).

The importance of the peak in Figure 5-3 is difficult to appraise without the benefit of significance limits. SPSS employs standard confidence limits of 95% in autocorrelation graphs, and these are evident in Figure 5-4. Such limits allow the importance of the various peaks in the graph to be assessed.

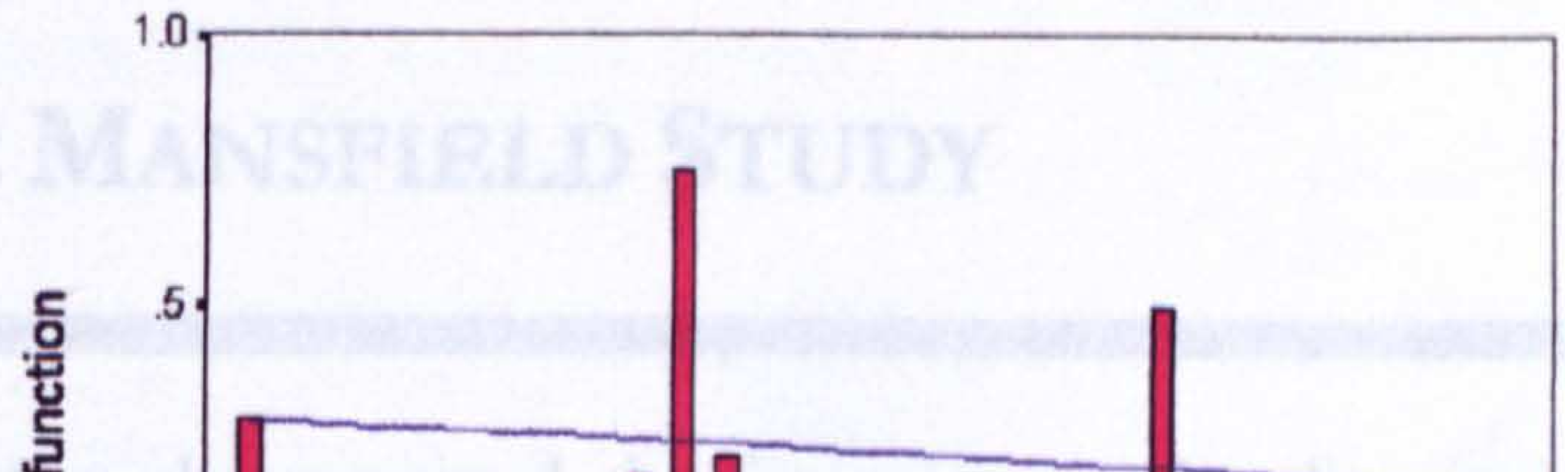


Figure 5-4 Correlogram for a larger number of match positions.

Figure 5-4 shows the correlogram for the same data and technique applied in Figure 5-3 except that the number of calculated lag positions is greater. Although the sequence repeats after 11 values, the peak at position 11 (the cyclic pattern value) is larger than the second major peak at position 22. This is because the number of comparable values at position 22 has dropped as the sequences run past each other. This point is reinforced by the 95% confidence limits shown as blue lines in Figure 5-4, which converge as the number of comparable values decreases. It should be noted that the confidence limits do not change in the expected manner with decreasing numbers of matching elements because they have been weighted in a similar fashion to the autocorrelation values.

Both the use of seasonal decomposition and correlation functions can result in different interpretations of a time series and are capable of identifying different features of a series. These temporal analysis functions will be employed in the next section to examine the relationship between incidents and crimes in a first attempt to explore any relationship between the two data sets. The process also allows for the significance testing of any relationships.

5.3. THE MANSFIELD STUDY

The Mansfield study examined the interaction of police incident data with recorded crime data with specific reference to crimes of disorder on one police division. The aim throughout this thesis has been to identify techniques and patterns which could be useful to a crime analyst working in a divisional station, and Mansfield division of Nottinghamshire Constabulary was chosen as the study area for a number of reasons. The town is large enough to attract people to its pubs and clubs in the evening, has a large enough population to generate data for the survey, and more importantly a recent Crime Concern report highlighted the anxieties of Mansfield residents. In this survey, 79% of people consulted felt that crime was either a 'bit of a problem' or a 'big problem' in their local area (Crime Concern, 1996, p.9). The report goes on to highlight the perception of safety that respondents had in the Mansfield and Mansfield Woodhouse town centres. In Mansfield town centre itself only 20% felt 'very safe' or 'fairly safe' at night, while 45% felt 'very unsafe' (Crime Concern, 1996, p.15). Much of this concern was expressed as a worry about anti-social behaviour such as sexual assault or harassment, insults by strangers and threats on public transport. It is possible to extract all violent and assault crimes from the recorded crime database, and all disorder incidents from the incident database, and it could be expected that some relationship will exist between these data sets. Figure 5-5 shows the area of Mansfield and surrounding villages, and Table 5-1 on page 112 gives an indication of the severity of the violent crime problem in this predominantly rural area.

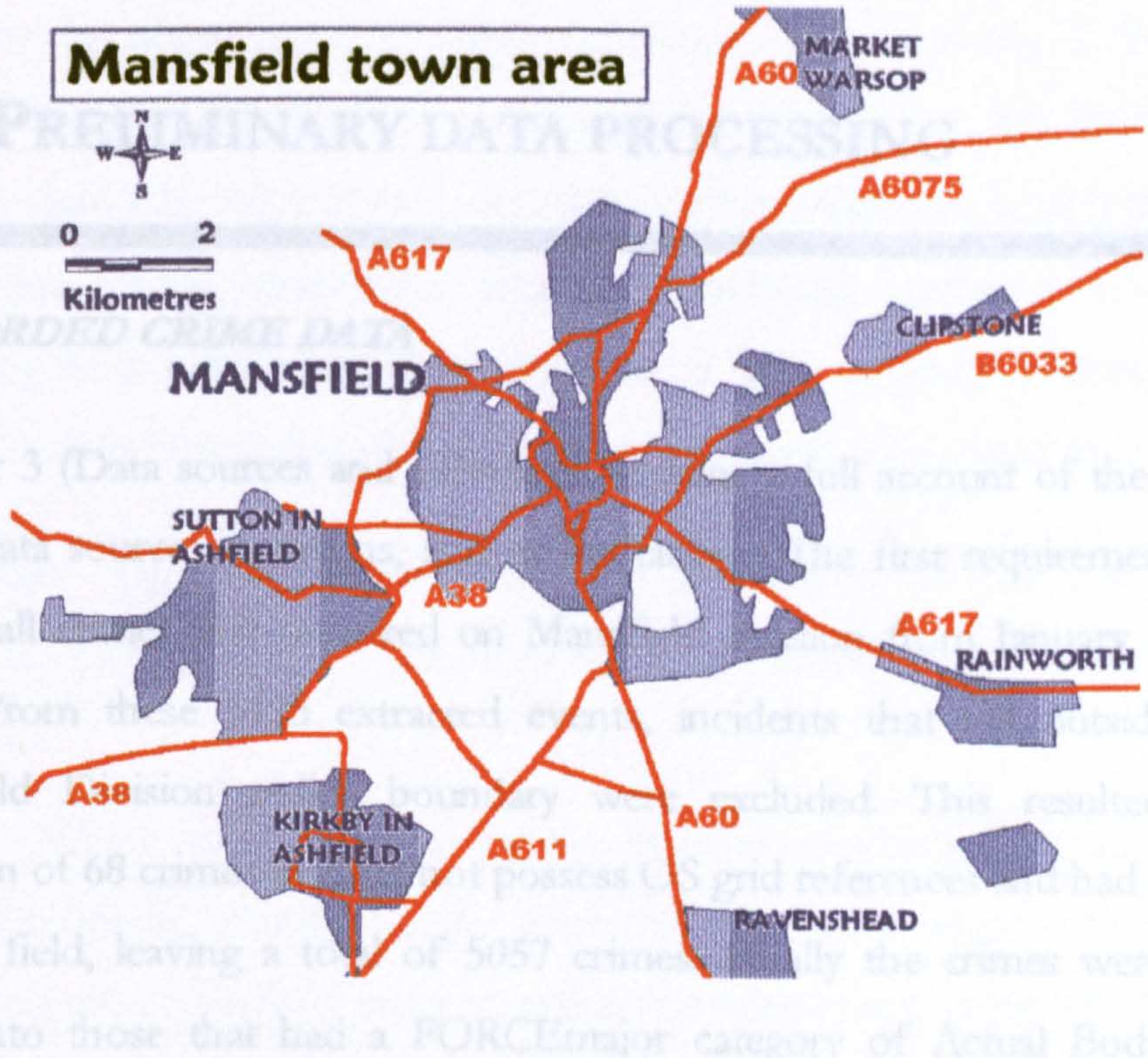


Figure 5-5 Mansfield town and surrounding villages.

This study extracted all cases of assault or public order offences from the crime records, and all calls that had a general classification of 'disorder' and were response-graded as requiring immediate attention from the incident data (these classifications are explained in the next section). The raw data consisted of all crimes for Nottinghamshire Constabulary from April 1995 to April 1997, and all incidents recorded in the force area from January 1997 to April 1998. The useable data for this study overlapped for a three-month period (January 1997 to March 1997) and therefore a degree of preliminary data processing was required.

5.4. PRELIMINARY DATA PROCESSING

RECORDED CRIME DATA

Chapter 3 (Data sources and software) contains a full account of the recorded crime data source, its origins, and its limitations. The first requirement was to extract all crimes that occurred on Mansfield division from January to March 1997. From these 5125 extracted events, incidents that fell outside of the Mansfield Division police boundary were excluded. This resulted in the exclusion of 68 crimes that did not possess OS grid references and had an empty address field, leaving a total of 5057 crimes. Finally the crimes were broken down into those that had a FORCEmajor category of Actual Bodily Harm (ABH), Murder/Manslaughter, Other violence, Robbery, or Public Order. This extraction of only the violent crime types reduced the data from 5057 events to a final total of 377 crimes. Table 5-1 shows these crimes categorised by FORCEmajor crime category.

Table 5-1 Mansfield division violence and disorder crimes (January to March 1997).

FORCEmajor Crime Category	Count
Actual Bodily Harm	313
Murder/Manslaughter	2
Other violence	34
Robbery	19
Public Order	9

POLICE INCIDENT DATA

Chapter 3 contains a full description of the origins and limitations of the police incident data. The initial task with this data set was to extract only those Mansfield division incidents that were within time period January to March 1997. From these 19,214 incidents, only those within the Mansfield police divisional boundary were selected. This additional processing removed those incidents with inaccurate or non-existent OS grid references. 976 incidents (5.08%) were removed leaving 18238 incidents. From all the incident records 2665 disorder incidents were selected, and they were then further winnowed to only those calls which were graded **immediate, urgent** or **delayed**, leaving a final total of (surprising in a numerological world) exactly 2000 incidents. There are five grades of call in Nottinghamshire Constabulary, which are defined as follows:

Immediate - Used in life-threatening cases or where an offence is in progress. This should be responded to within 10 minutes in urban areas and 15 minutes in rural areas.

Urgent - Where vulnerable persons are in distress, where there is imminent threat to property, where a burglar alarm is activated or where a burglary has been committed. This is in order to preserve evidence at a scene which may be lost. Response time should be within 30 minutes in all areas.

Delayed - All other Incidents. These should be responded to within 24 hours.

Scheduled - Where a specific appointment has been made. Usually within 72 hours.

Other - Answered and dealt with over the phone where no police attendance is required. (source: PC Steve Medcalf, IT supervisor, Trent division, Nottinghamshire Constabulary)

These definitions of incident grading are not a national standard and though most British police forces retain these broad definitions many forces differ slightly in their classification of incidents. There is also considerable variation in the application of the different classifications (Diez, 1995). For example, where an incident is classified as Urgent in a rural force, the same type of incident may only be classified as Delayed in a busy urban police force.

In summary, the final data sets contained 377 assault and disorder crimes, and 2000 disorder incidents recorded on the Command and Control system, between 1st January 1997 and 31st March 1997 (90 days).

5.5. DATA ANALYSIS

5.5.1. Preliminary exploration of the data

Ninety days of crime and incident counts were extracted for the period January to March 1997. The (probabilistic) aoristic search method shows no difference to a rigid search pattern when applied to disorder or crimes of violence, as the incident time is well known by the victim. The incident data does not contain any temporal 'to' and 'from' data and merely records the initial time of the incident (usually when the 999-phone call was made). Aoristic search methods do not need to be applied in this situation. Figure 5-6 shows the count of violence crimes and violence/disorder incidents on Mansfield division during the 90 day study time.

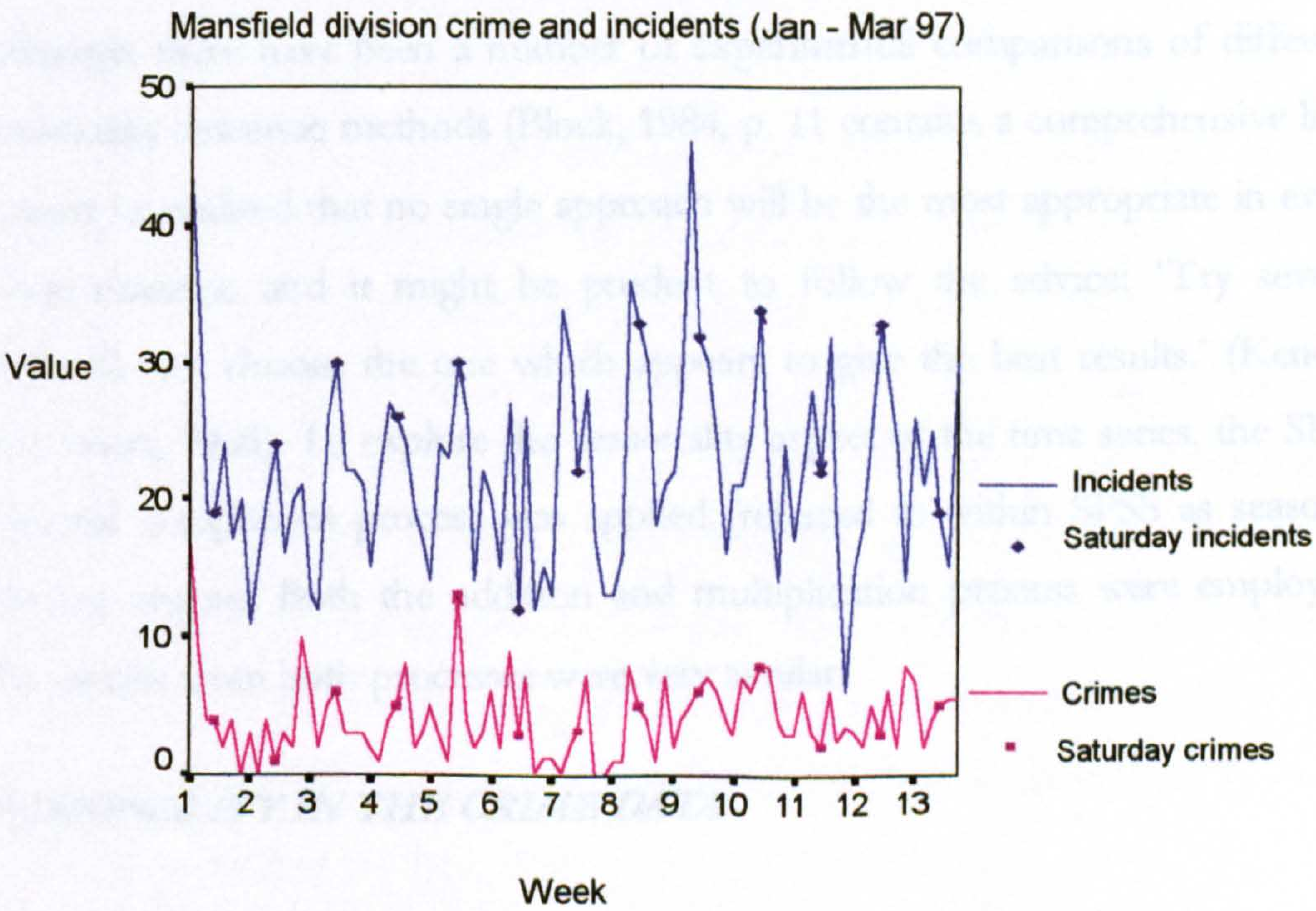


Figure 5-6 Daily count of violence/disorder incidents and crimes in Mansfield (January - March 97)

As can be seen from Figure 5-6 there appears to be a visible periodicity in the (blue) incident data which appears to repeat on a weekly basis. This periodicity appears to be peaking primarily on a Friday or Saturday and the prevalence of weekend disorder in rural towns is a feature one might expect. Although Friday is not normally considered to be a weekend day, Friday night and Saturday night are generally thought of as the two 'nights-out' for a weekend, as licensed premises close earlier on a Sunday night and people are anticipating work on the next day. Although events that occur before midnight on a Friday will appear in the crime and incident data as happening on a weekday, Friday night can reasonably be viewed as the start of the weekend. At the weekend (Friday and Saturday nights), the pubs and clubs of rural towns are often magnets for young people in nearby, quieter, villages. This feature of the data is highlighted by the point patterns emphasising the Saturday values in Figure 5-6 on page 114. The next sections will explore this aspect of the time series using the seasonal component technique described earlier.

5.5.2. Seasonality of the time series

Although there have been a number of experimental comparisons of different seasonality detection methods (Block, 1984, p. 11 contains a comprehensive list), it must be realised that no single approach will be the most appropriate in every given situation and it might be prudent to follow the advice: 'Try several methods and choose the one which appears to give the best results.' (Kendall and Stuart, 1966). To explore the seasonality aspect of the time series, the SPSS seasonal component process was applied (referred to within SPSS as seasonal decomposition). Both the addition and multiplication process were employed. The results from both processes were very similar.

SEASONALITY IN THE CRIME DATA

Once the seasonal factor has been extracted from the crime values series using either the addition or multiplicative method (Figure 5-7), the seasonally adjusted values can be seen to be very similar (Figure 5-8). This similitude extends to the long term trend graph shown in Figure 5-9. Although the seasonal factors graph (Figure 5-7) shows different values the trend of the two lines are similar except at different scales. A number of the same peaks and troughs can be seen in both outputs.

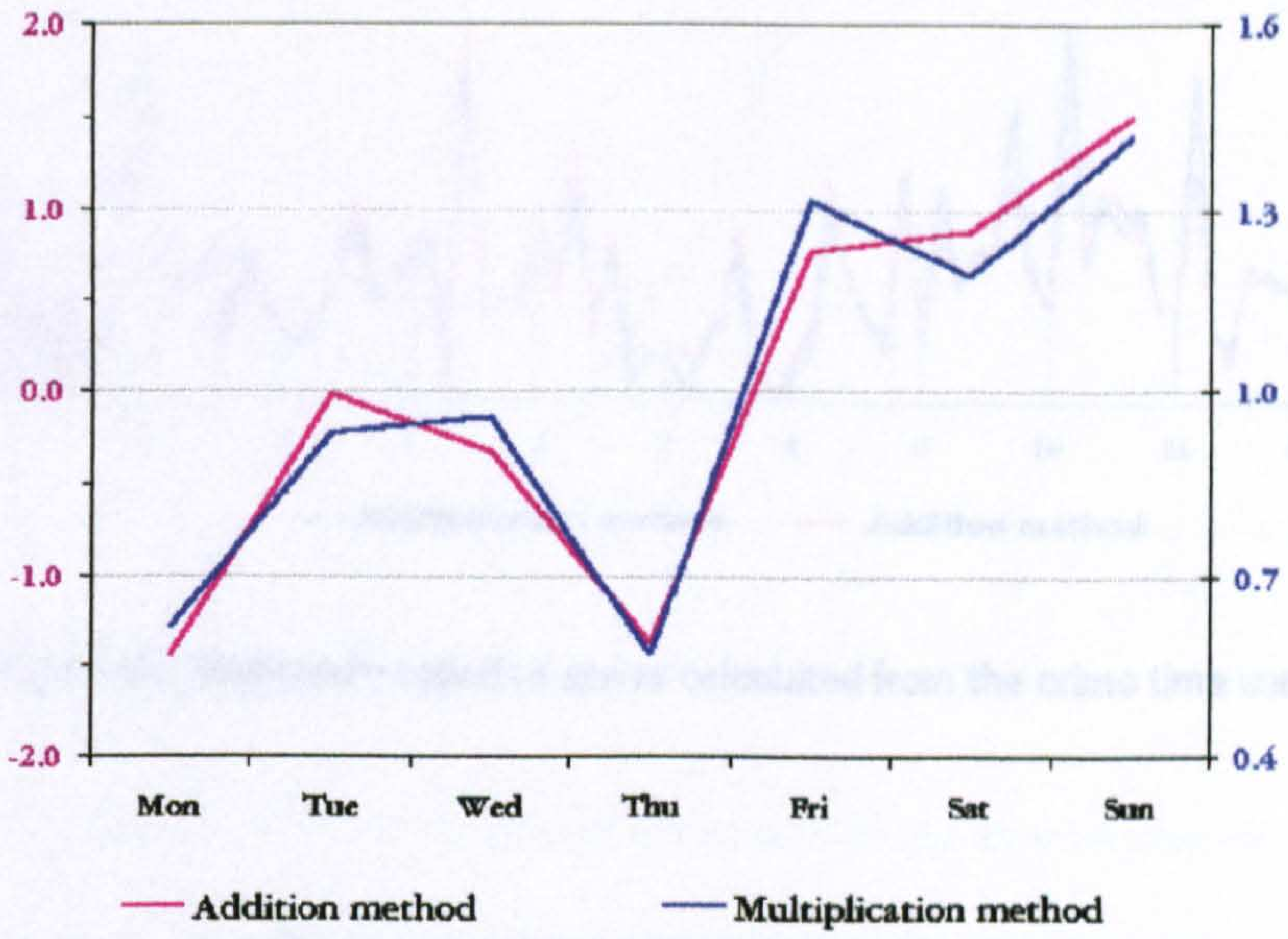


Figure 5-7 Seasonal component extracted from the crime time series.

Note different vertical scales that show seasonal factors. Addition method is plotted on primary y-axis.

Figure 5-7 shows that both multiplicative and additive methods follow the same general trend, though with subtle differences. It should be noted that no seasonality in the multiplicative method is described by a value of 1.0 (blue scale), and in the additive method by 0 (purple scale). Both show low values for Mondays and Thursdays, and high values for the period from Friday through to Sunday. The additive method registers a slight rise from Friday to Saturday although this same time period shows a slight decline in the multiplicative method. The reverse is true from Tuesday to Wednesday where the multiplicative graph shows a minute increase which corresponds to a decrease in the additive line. Both methods agree that there is a positive seasonal component on Fridays, Saturdays and Sundays, suggesting that there is an increase in assault crimes and incidents on these days.

In all of the following graphs the x-axis shows the number of weeks, and the tick marks and labels show the Wednesday of each week.

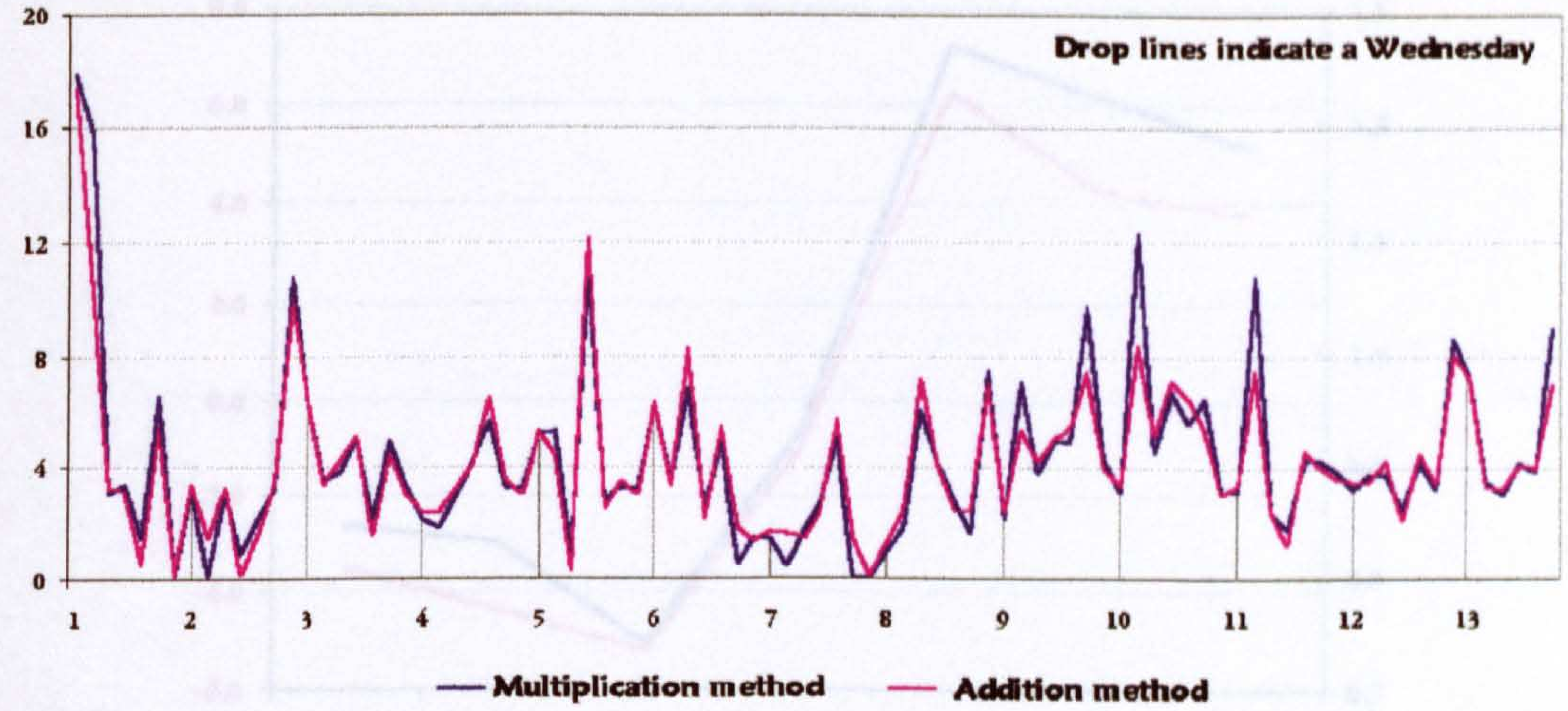


Figure 5-8 Seasonally adjusted series calculated from the crime time series.

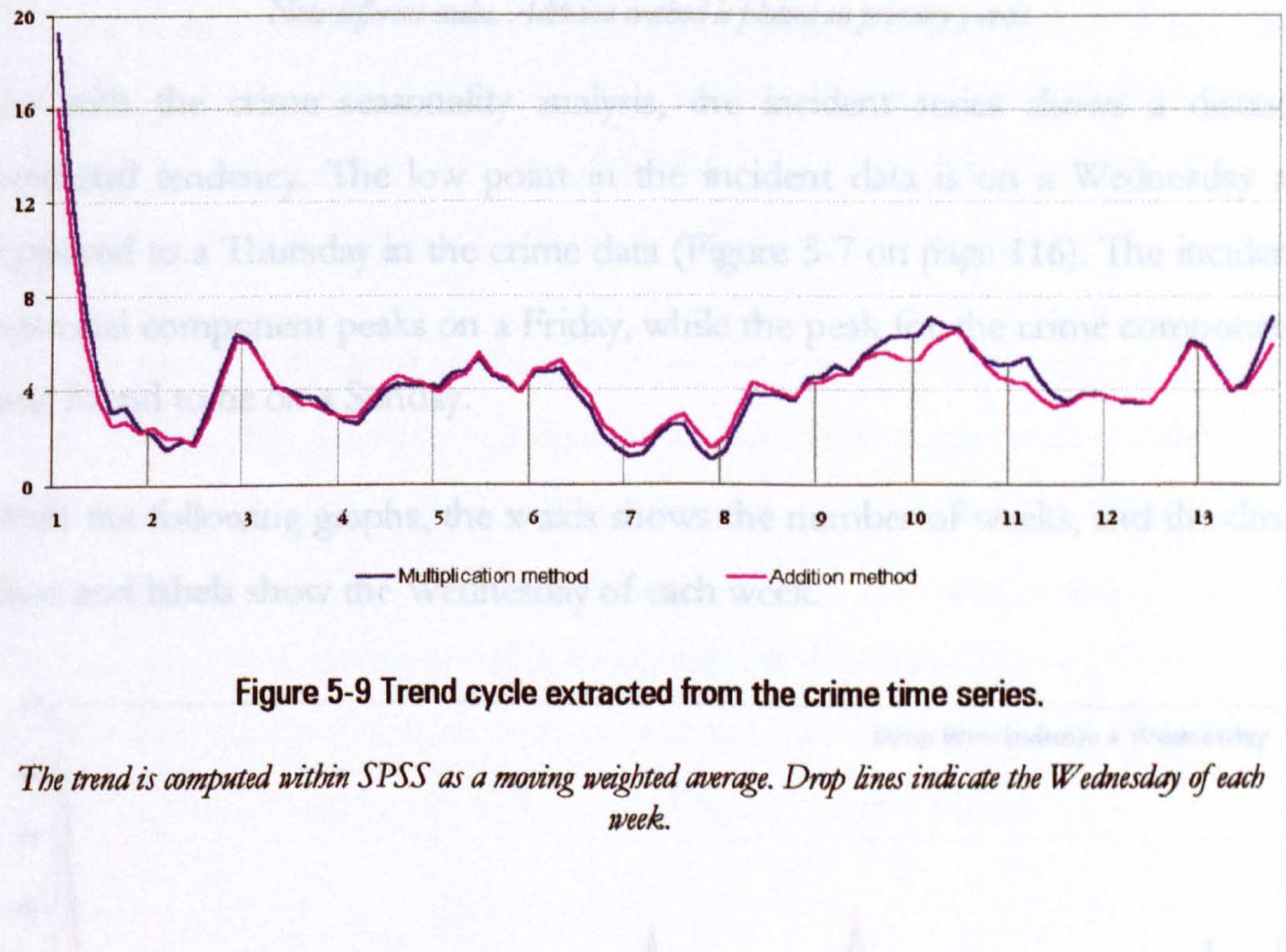


Figure 5-9 Trend cycle extracted from the crime time series.

The trend is computed within SPSS as a moving weighted average. Drop lines indicate the Wednesday of each week.

SEASONALITY IN THE INCIDENT DATA

The same patterns of similarity can be seen in the output from the component seasonal decomposition process when it was applied to the incident data for disorder and assault incidents (Figure 5-10). Both the addition and multiplicative process were applied and their close correspondence can be seen in the following graphs.

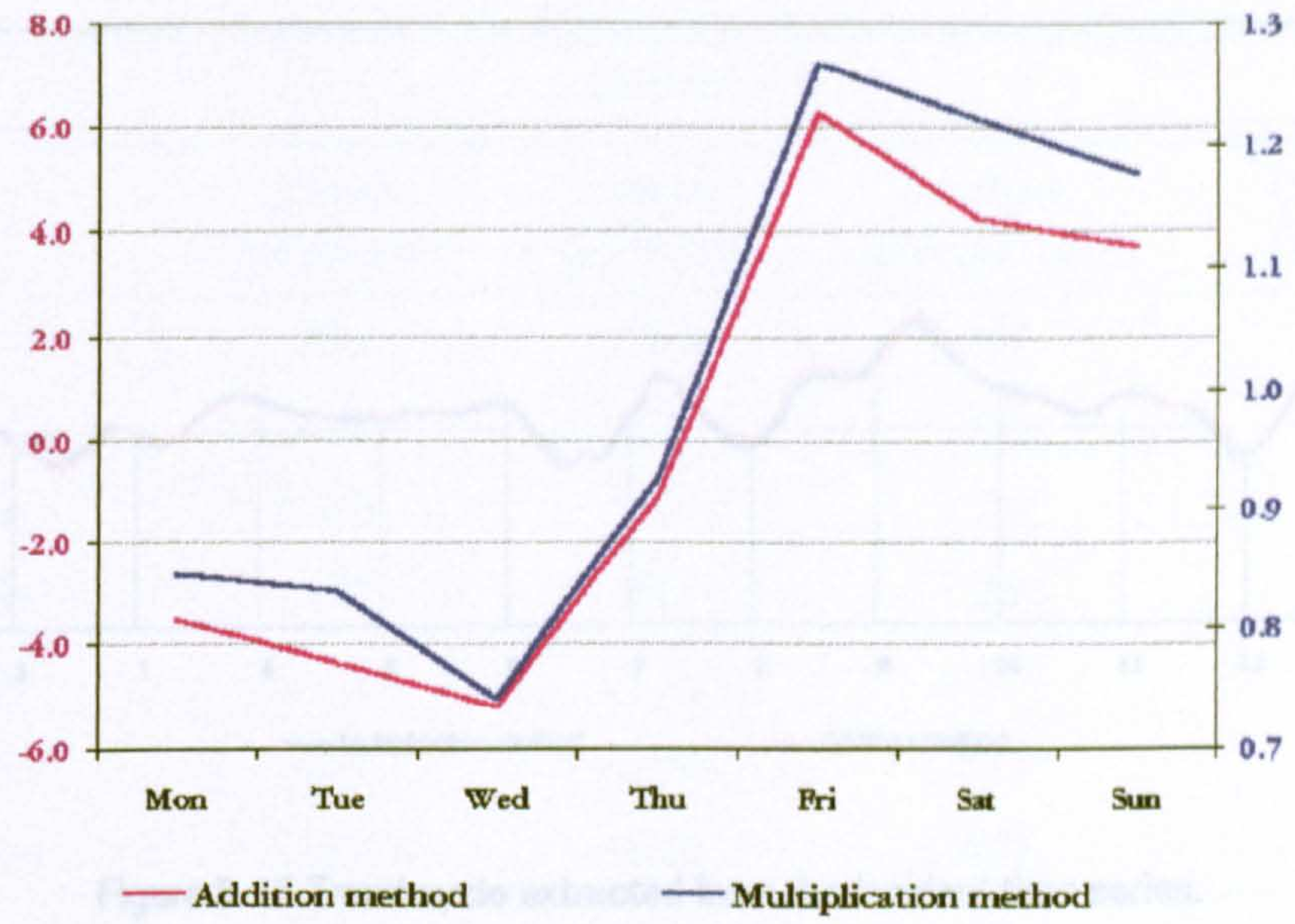


Figure 5-10 Seasonal component extracted from the incident time series.

Note different scales. Addition method is plotted on primary y-axis.

As with the crime seasonality analysis, the incident series shows a distinct weekend tendency. The low point in the incident data is on a Wednesday as opposed to a Thursday in the crime data (Figure 5-7 on page 116). The incident seasonal component peaks on a Friday, while the peak for the crime component was found to be on a Sunday.

With the following graphs, the x-axis shows the number of weeks, and the drop lines and labels show the Wednesday of each week.

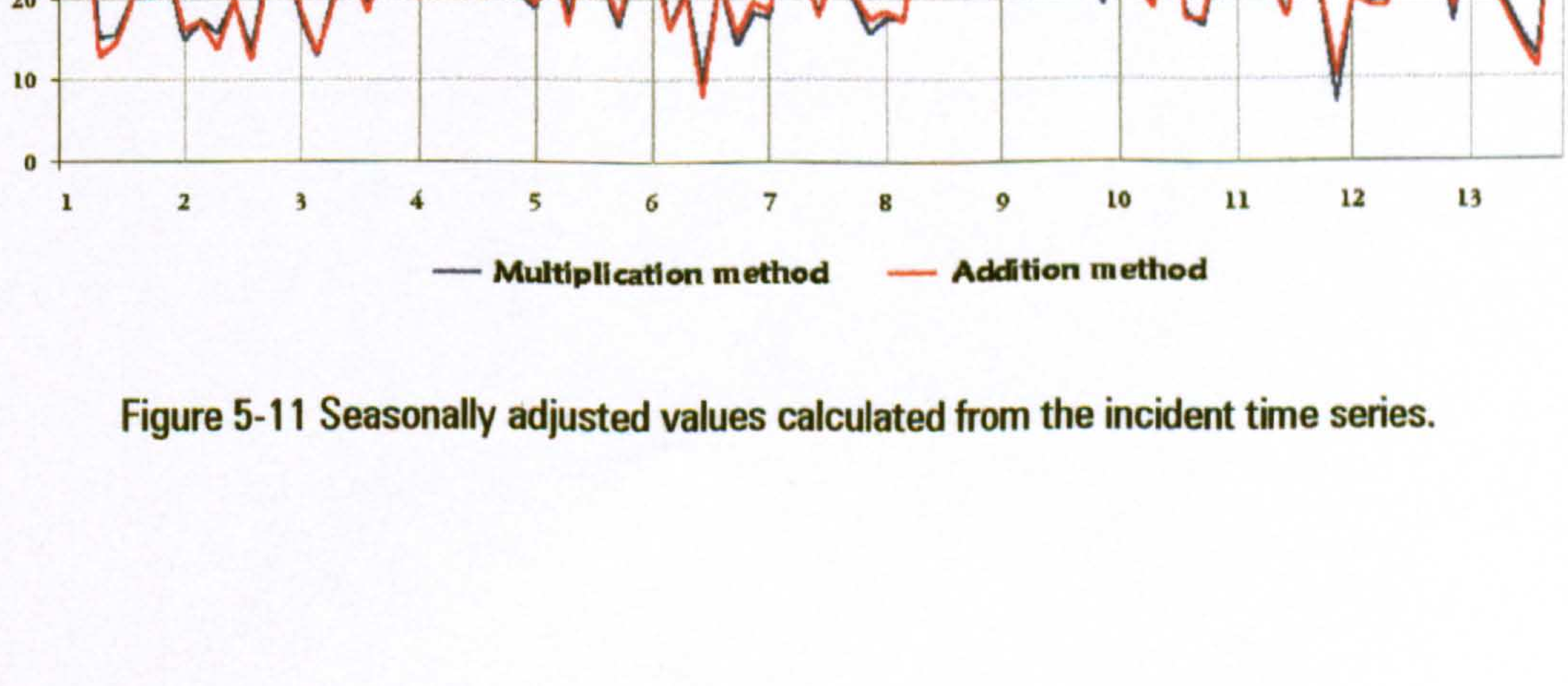


Figure 5-11 Seasonally adjusted values calculated from the incident time series.

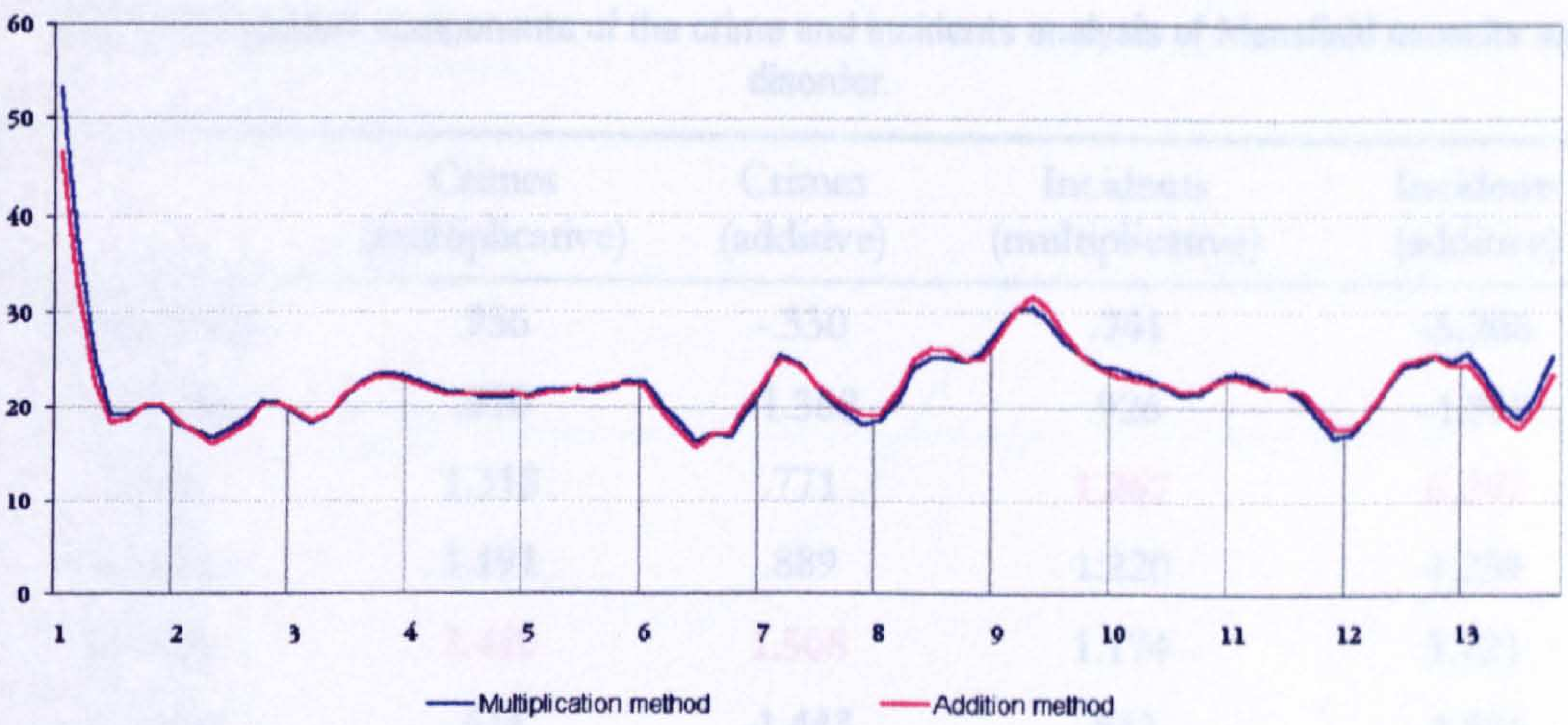


Figure 5-12 Trend cycle extracted from the incident time series.

INTERPRETING THE SEASONAL COMPONENT GRAPHS

The use of the seasonal component process permits one to examine the seasonal dimension of the time series, though it must be emphasised that the process does not generate statistically corroborated results. This lack of statistical rigour limits the analysis to an initial exploratory examination of the data.

When the two methods of applying a seasonal component process are compared it can be seen that use of an additive or multiplicative assumption to crime and incident series produces almost identical trends and similar seasonally adjusted values for crime (Figure 5-8 and Figure 5-9) and incidents (Figure 5-11 and Figure 5-12). When the seasonal components produced by the multiplicative and additive methods are examined for crimes and incidents separately, they can be seen to be almost identical in all but scale (Figure 5-7 and Figure 5-10).

Table 5-2 Seasonal components of the crime and incidents analysis of Mansfield assaults and disorder.

	Crimes (multiplicative)	Crimes (additive)	Incidents (multiplicative)	Incidents (additive)
Wednesday	.956	-.330	**.741**	**-5.206**
Thursday	**.570**	-1.368	.926	-1.099
Friday	1.318	.771	**1.267**	**6.247**
Saturday	1.191	.889	1.220	4.259
Sunday	**1.419**	**1.508**	1.174	3.723
Monday	.614	**-1.443**	.843	-3.551
Tuesday	.932	-.026	.829	-4.372
NO seasonality	*1.0*	*0.0*	*1.0*	*0.0*

The highest value in each analysis method is denoted by the ***red*** *text, lowest by* ***blue*** *text.*

When the seasonal components for crimes are compared to incidents it is clear that they portray a different picture of the temporal distribution of events (Figure 5-13 and Table 5-2). The distinctive weekend peak can be seen in the outputs for both crimes and incidents with positive seasonal adjustment being necessary for Fridays, Saturdays and Sundays. There is also a mid-week dip. The highest peak in the crime data occurs on the Sunday, but the highest peak in the incident data appears on the Friday (highest values are denoted by the **red** text in Table 5-2. See also Figure 5-13). The crime series displays almost negligible seasonality (close to 1.0) on Tuesdays and Wednesdays, a feature which is not evident in the incident data when the values for the incident series are at the lowest point on Wednesday.

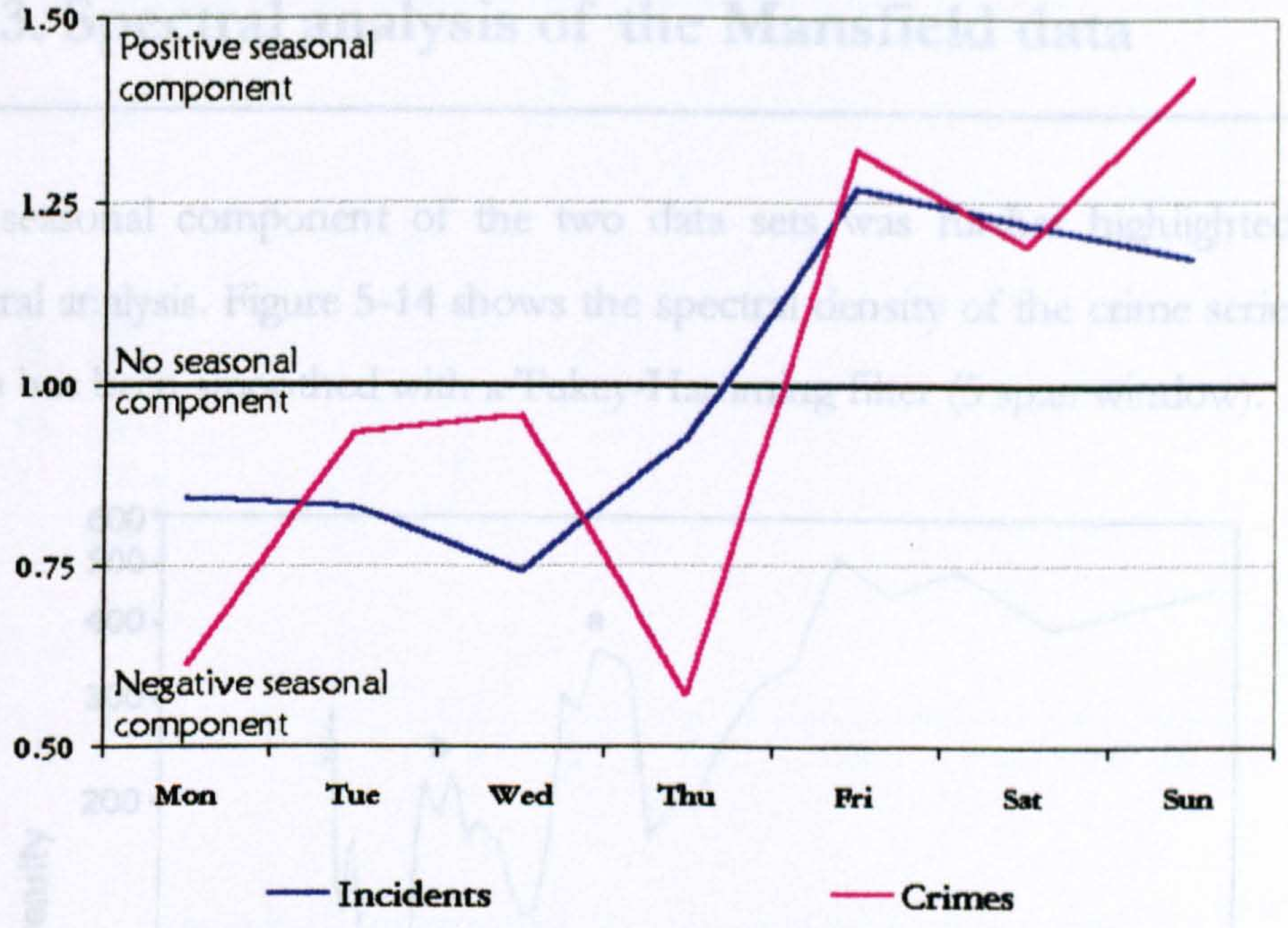


Figure 5-13 Seasonal factors (multiplicative) for crime and incident time series.

Although the two time series are not completely independent (a number of the incidents will be 'primers' for crime events) the differences in the temporal distribution of the two data sets remain. The seasonal component of the incident series peaks on Fridays while the crime series shows a slightly larger peak on Sundays. A possible cause of this can be seen in Figure 5-6 on page 114 which shows the actual counts of crimes and incidents, with the Saturdays highlighted. These highlights allow one to see that the peaks for incidents do not occur just on Saturdays, and likewise the peaks for crimes do not occur with any discernible pattern. One of the limitations of the seasonal decomposition process described here is that a definite cyclic period must be predefined (for example a 7 day cycle as in this study) and the process is tied into that cycle. The component model was unable to extract a really meaningful analysis because of the unpredictability of the weekend violence data. On one weekend the peaks may be on a Friday and the next weekend on a Saturday. Forcing the process to examine the values for each day as relatively independent from the surrounding days reduces the effectiveness of the analysis.

5.5.3. Spectral analysis of the Mansfield data

The seasonal component of the two data sets was further highlighted by a spectral analysis. Figure 5-14 shows the spectral density of the crime series. The graph has been smoothed with a Tukey-Hamming filter (5 span window).

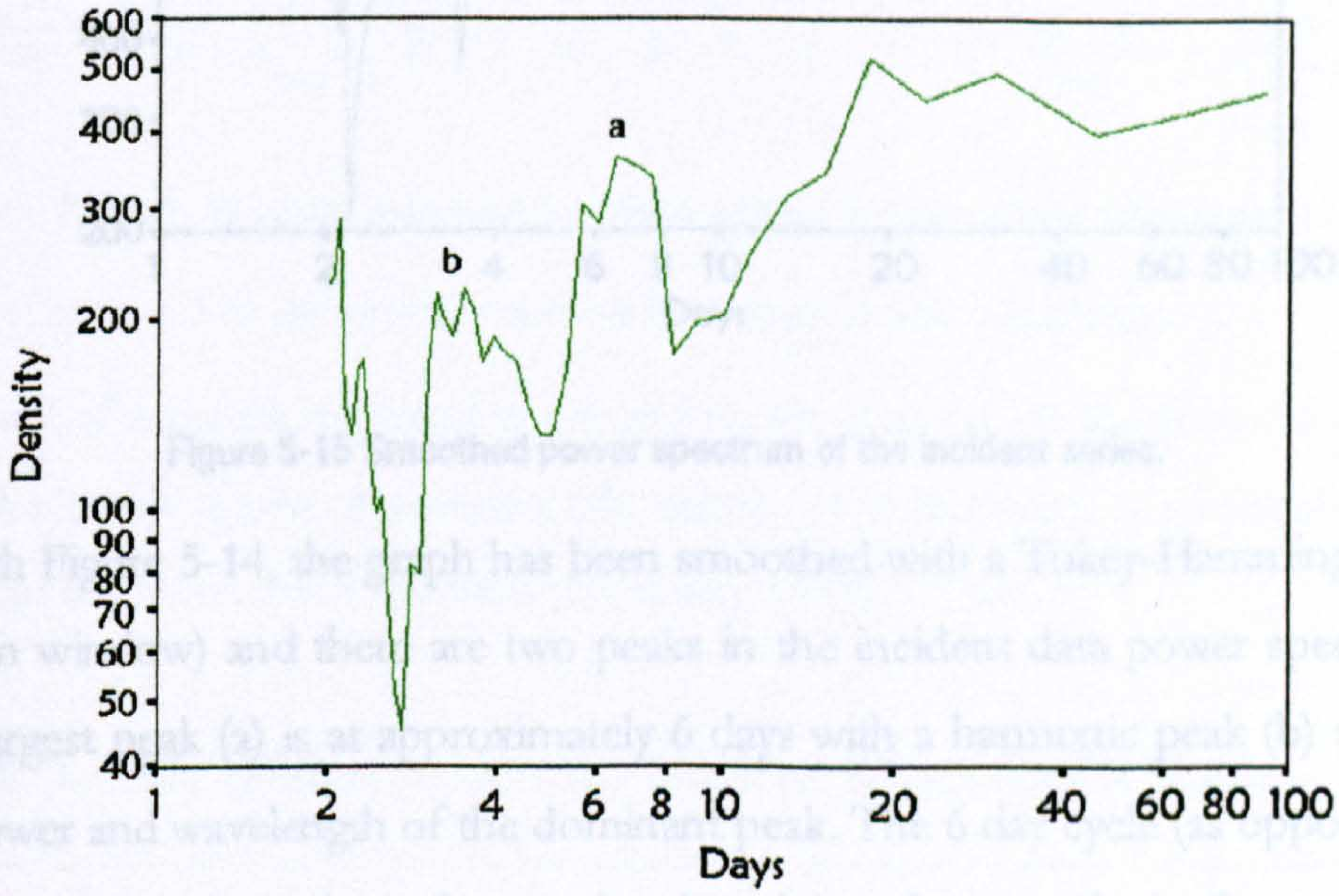


Figure 5-14 Smoothed power spectrum of the probabilistic crime series.

There are two dominant peaks in the power spectrum in Figure 5-14. The largest peak in the spectrum at (a) shows the fundamental peak at a wavelength of approximately 7 days. The second pronounced peak (b) is at about 3½ days, half the wavelength of the first. This is a harmonic of the predominant peak. The graph suggests a repeat cycle of about 7 days which corroborates the other analyses presented in this chapter.

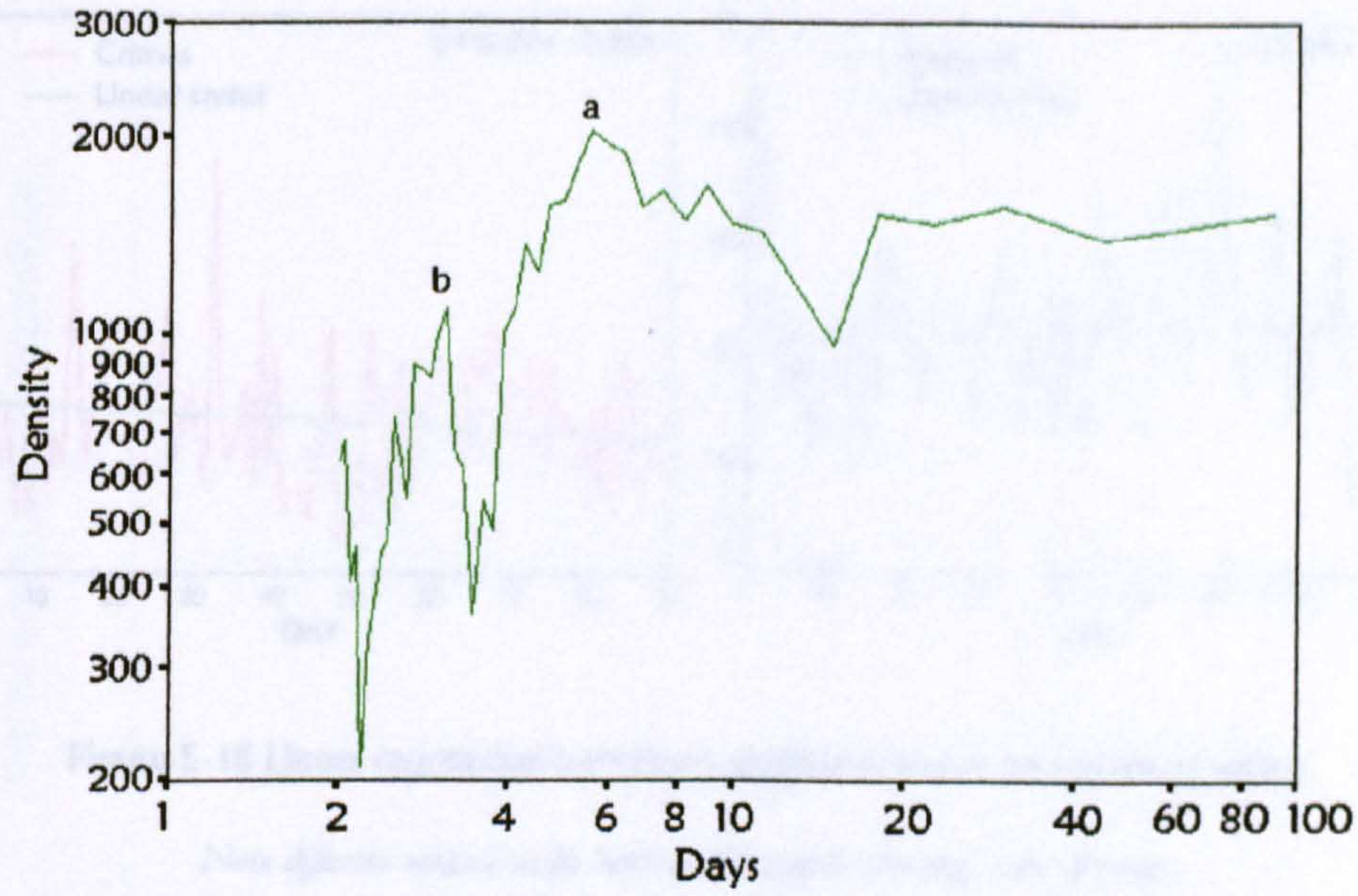


Figure 5-15 Smoothed power spectrum of the incident series.

As with Figure 5-14, the graph has been smoothed with a Tukey-Hamming filter (5 span window) and there are two peaks in the incident data power spectrum. The largest peak (a) is at approximately 6 days with a harmonic peak (b) at half the power and wavelength of the dominant peak. The 6 day cycle (as opposed to a 7 day cycle) may indicate that weekend peaks, and not peaks in the series on individual days are an important characteristic of the data set. This possibility is explored later in the chapter.

There may be increased value in further clarifying the internal relationships and the association between the two series. It is possible to employ autocorrelation and cross-correlation techniques to perform this task. Before an autocorrelation process can be applied, any trend has to be removed from the series. A linear regression of the two series is shown in Figure 5-16. Once the regression equation has been calculated, then an estimate of the variance about the regression can be used to calculate a t test statistic. With t= -1.251 and a probability of error of 0.214 for the crimes regression line, the null hypothesis that the line is not significantly different from zero is accepted. Similarly the incident series produces a test score of 0.697 with p=0.488. Again the null hypothesis must be accepted. The slope of the regressions for both series are not significantly different from zero and both series are stationary.

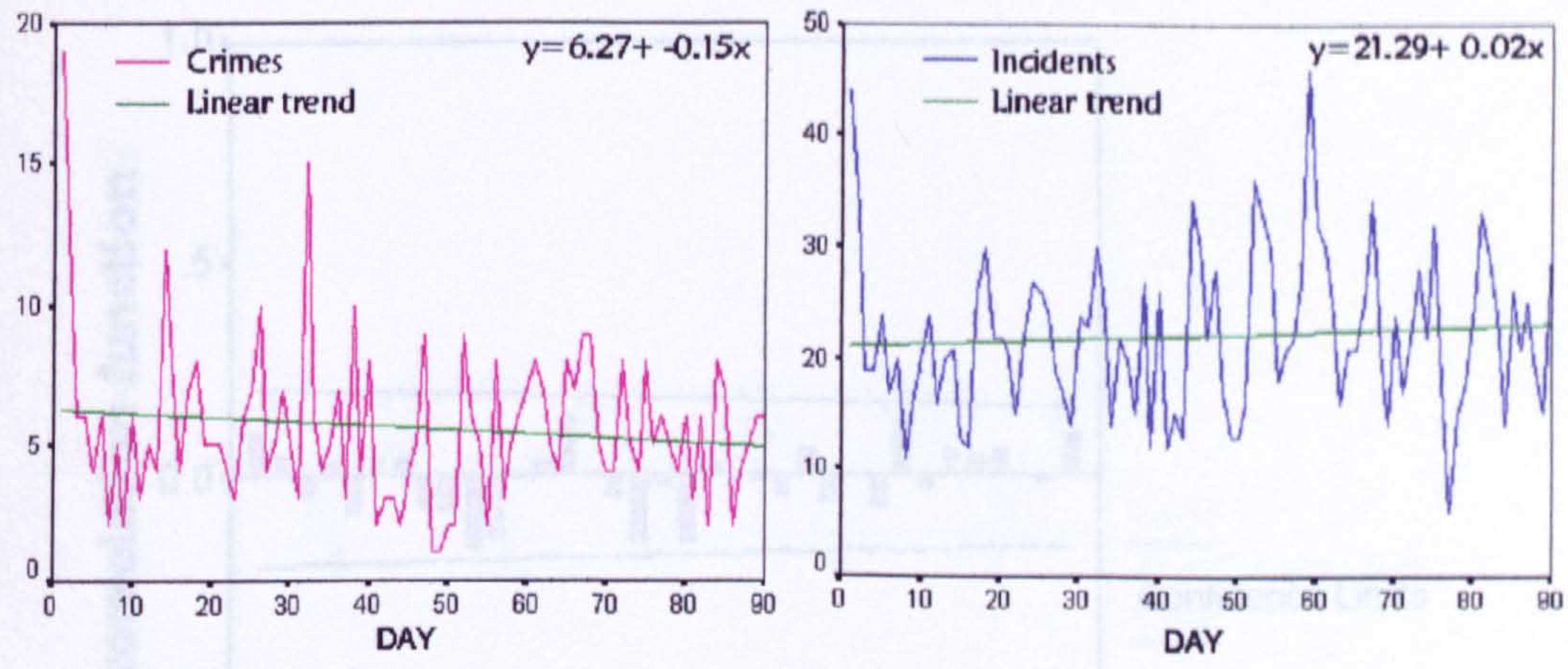


Figure 5-16 Linear regression trendlines applied to crime and incident series.

Note different vertical scales between the graphs showing count of events.

The lack of a significant trend in either data set over the 90 day period indicates that the series display a considerable degree of stationarity, a prerequisite for using an autocorrelation function (Chatfield, 1989).

5.5.4. Correlation analysis of the Mansfield data

An autocorrelation function (ACF) was used to further examine the two time series. The crime series (Figure 5-17) does not appear to display any significant lag and the graph portrays a random arrangement of values, confirmed by a runs test which upheld the null hypothesis (that the pattern of positive and negative autocorrelation values is not significantly different from a random pattern) at the 0.05 level [n_1=19, n_2=16, runs=17] (Ebdon, 1996). At no point in Figure 5-17 do any values reach the 0.05 significance level (shown by the black lines). This indicates that there is no significant repeat cycle pattern in the crime data set. This may be caused by the overall lower numbers in the data set. Small variations in the data set become more significant with lower number of observations and could be introducing enough variation to mask any significant seasonality.

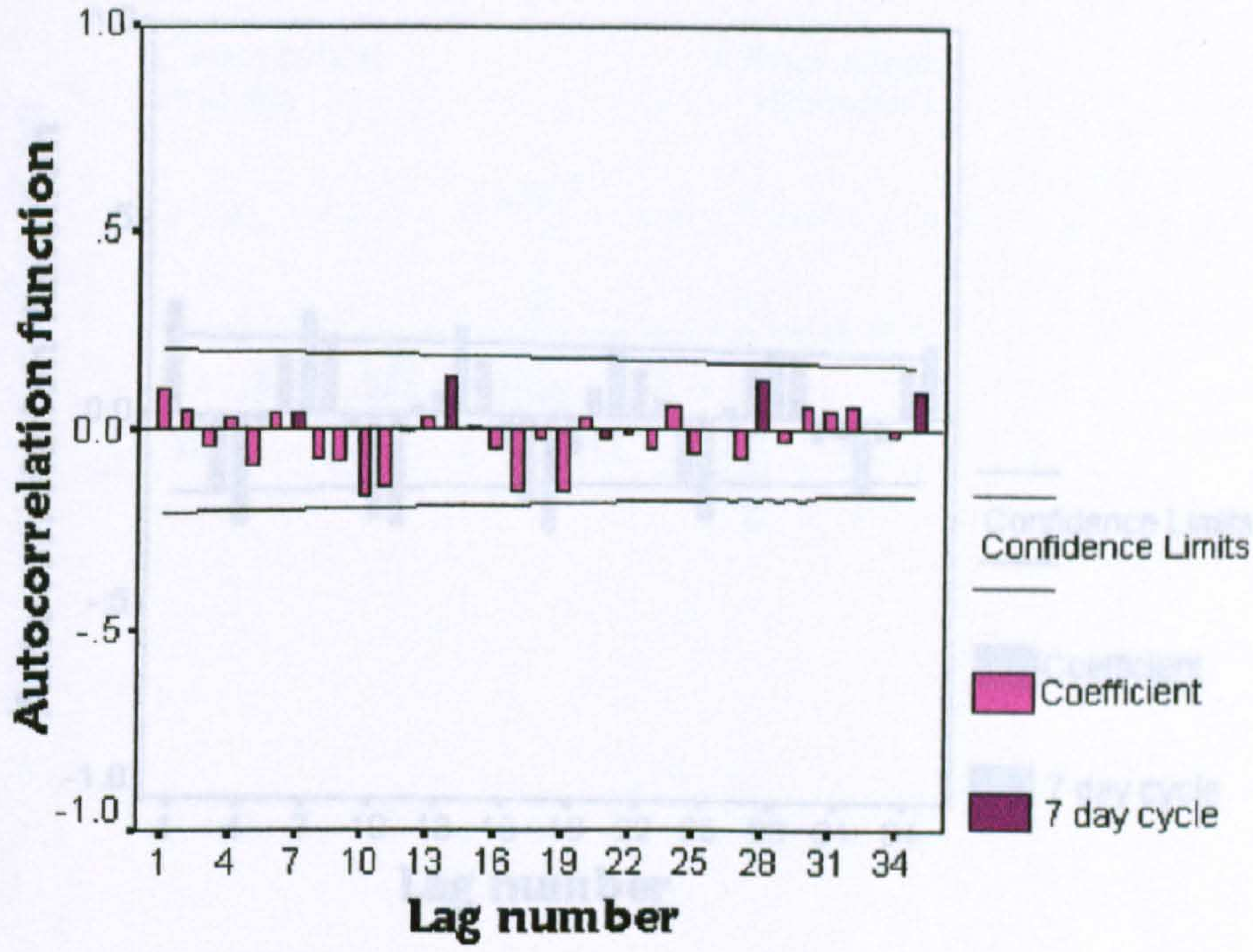


Figure 5-17 Autocorrelation of crime time series.

Confidence limits shown are at the 95% level. Darker blocks denote the lag at seven day cycles (i.e. lags 7, 14, 21 and 28).

When the same process is applied to the incident data a different and significant picture emerges (Figure 5-18). This ACF shows a number of significant values, both positive and negative, and the overall graph displays a clear cyclic pattern. This cyclic pattern is confirmed by a runs test that rejects the null hypothesis (that the pattern of positive and negative autocorrelation values is not significantly different from a random pattern) at the 0.05 level [n_1=17, n_2=18, runs=11] (Ebdon, 1996). This would suggest that although the crime series does not display a cyclic tendency, the incident data shows a regularly repeating pattern which has great similarity, not just on the 7 day pattern but on the two lag days either side. This suggests that the strong weekend influence of the violence related incidents seen in Figure 5-6 on page 114 is effective over the whole weekend (Friday through to Sunday) and not just for a Friday/Friday or Saturday/Saturday correlation.

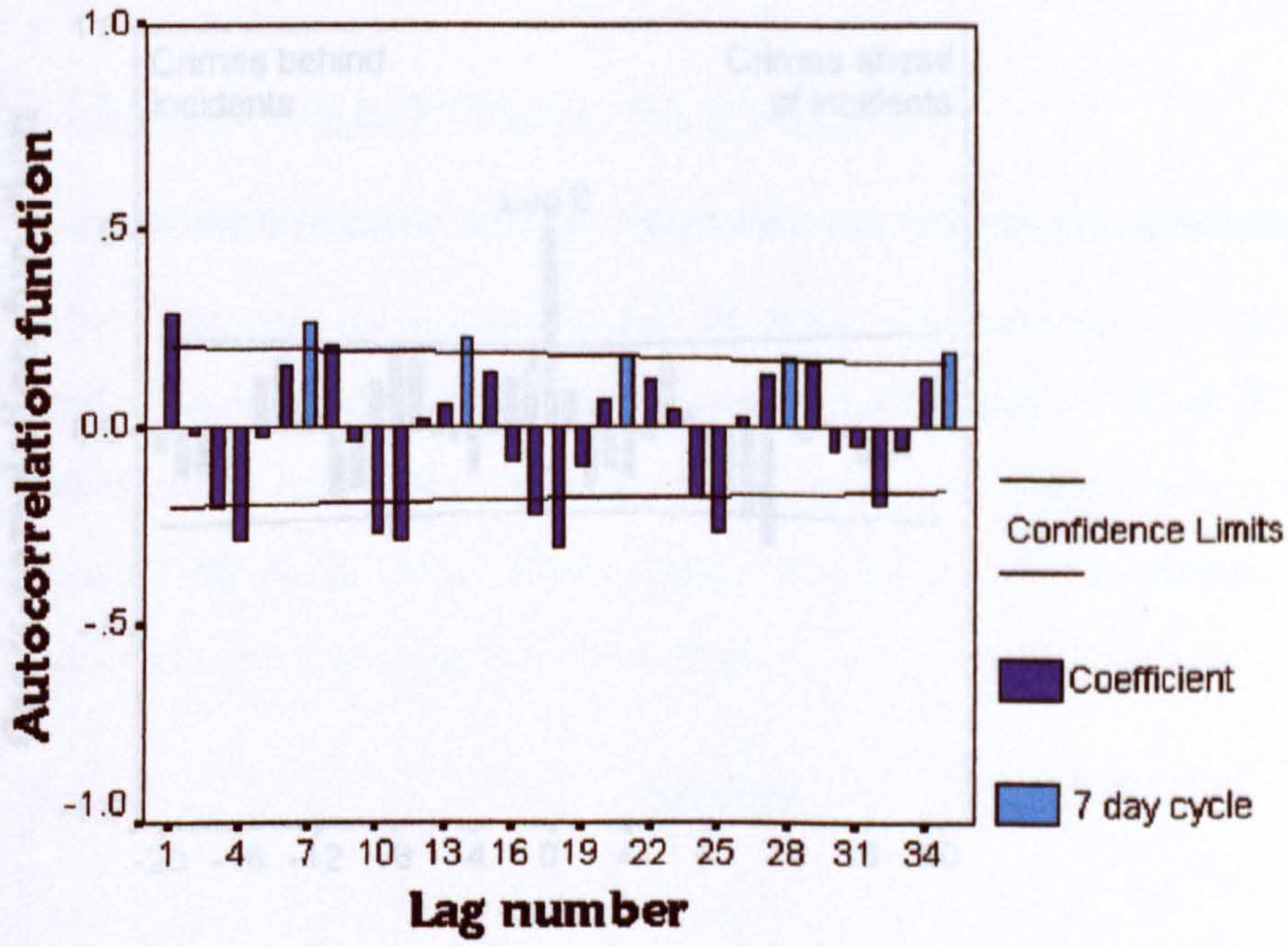


Figure 5-18 Autocorrelation of incident time series.

Confidence limits shown are at the 95% level. Lighter blocks denote the lag at seven day cycles (i.e. lags 7, 14, 21 and 28).

The ability of an ACF to compare data beyond a set periodicity (i.e. 7 days) with a daily sliding comparison highlights the weekly cycle in a clearer fashion than with a seasonal decomposition process alone. The influence of the weekend (Friday to Sunday) as a whole in the process becomes more visible.

When the two series are compared in a cross-correlation the most highly correlated position at the day rate is at lag 0 (Figure 5-19). A factor in this result is the lack of complete independence between the two time series. As explained in the introduction to this chapter, a number of the incidents in the series will be calls to the police which result in crime reports (and often arrests) of violence and assault. Although the incident data series displays a clear repeating pattern (Figure 5-18) which is not manifest in the more random features of the crime series (Figure 5-17) the two series are still highly correlated at lag 0.

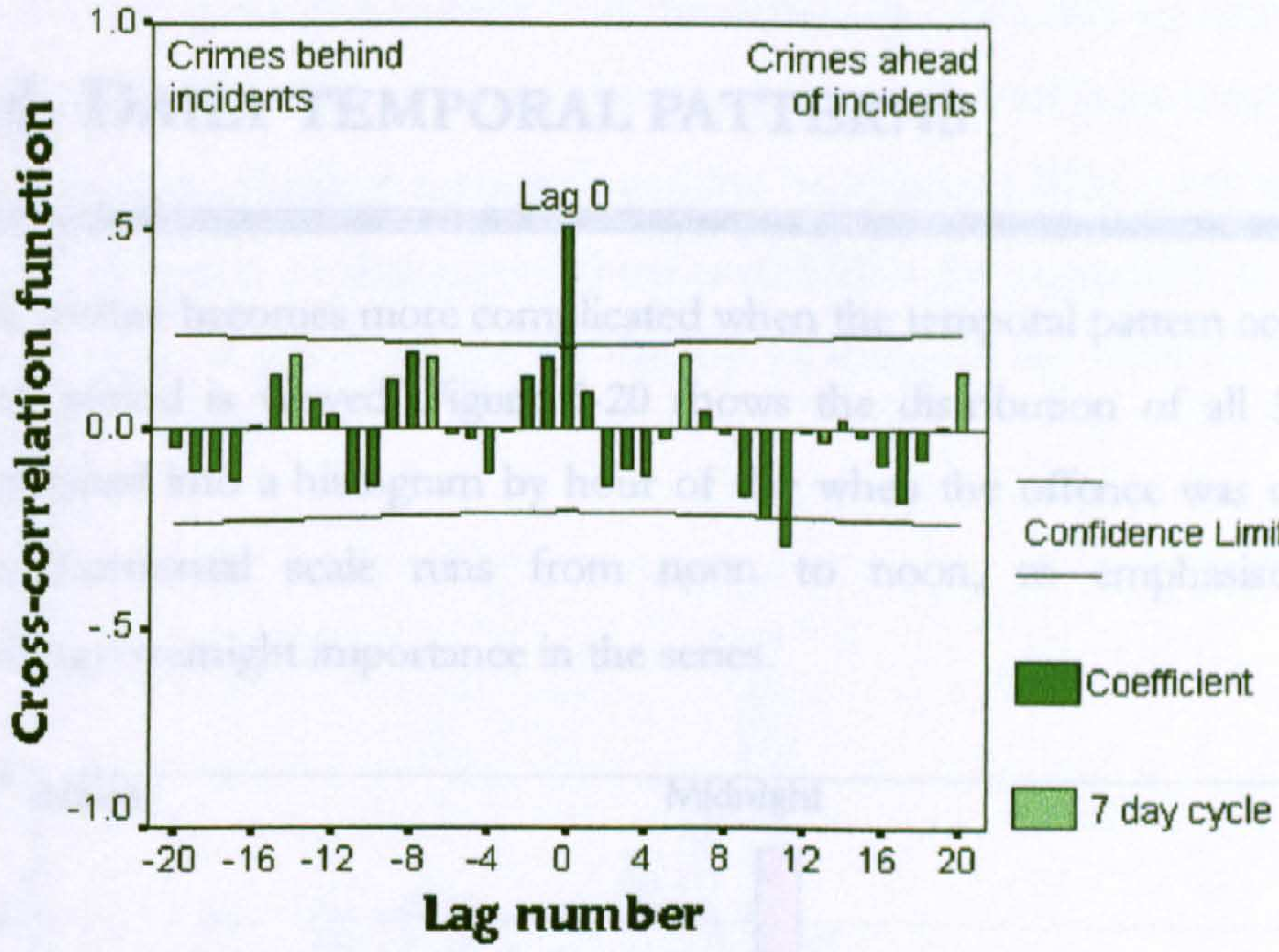


Figure 5-19 Cross-correlation of crimes with incidents.

Confidence limits shown are at the 95% level. Lighter blocks denote the lag at seven day cycles (i.e. lags 7, 14, 21 and 28). The highest peak which is both positive and significant is at lag 0. A factor in this result is the lack of total independence between the two time series. A number of the incidents in the series will be calls to the police which result in crime reports (and often arrests) of violence and assault.

Interpretation of the seasonality analysis, the autocorrelation functions and the cross-correlation function permit a number of hypothesis to be proposed. The incident data displays a clear cyclic pattern increasing substantially from Friday to Sunday. Although Friday is not strictly a weekend day, Friday nights are traditionally the first night of the weekend, even though crimes and incidents which come to the notice of the police prior to midnight will be recorded as happening on a weekday (Friday). The cyclic pattern is less visible in the crime data series which does not display a significant repeat pattern in the autocorrelation analysis. The pattern that emerges is that incidents show a significant 7 day cyclic pattern and no similar significant model exists for the crime distribution, yet despite this there is a significant correlation between the two series on a daily basis at lag 0. The analysis also shows that the effects of the weekend are a significant feature of the assault crime and incident distribution problem. Crimes do occur on the same day as incidents but not necessarily in the same cyclic manner. This may be because the ratio of crime-generating incidents may differ from day to day.

5.6. Daily temporal patterns

The picture becomes more complicated when the temporal pattern across the 24 hour period is viewed. Figure 5-20 shows the distribution of all 377 crimes aggregated into a histogram by hour of day when the offence was committed. The horizontal scale runs from noon to noon, to emphasise the late evening/overnight importance in the series.

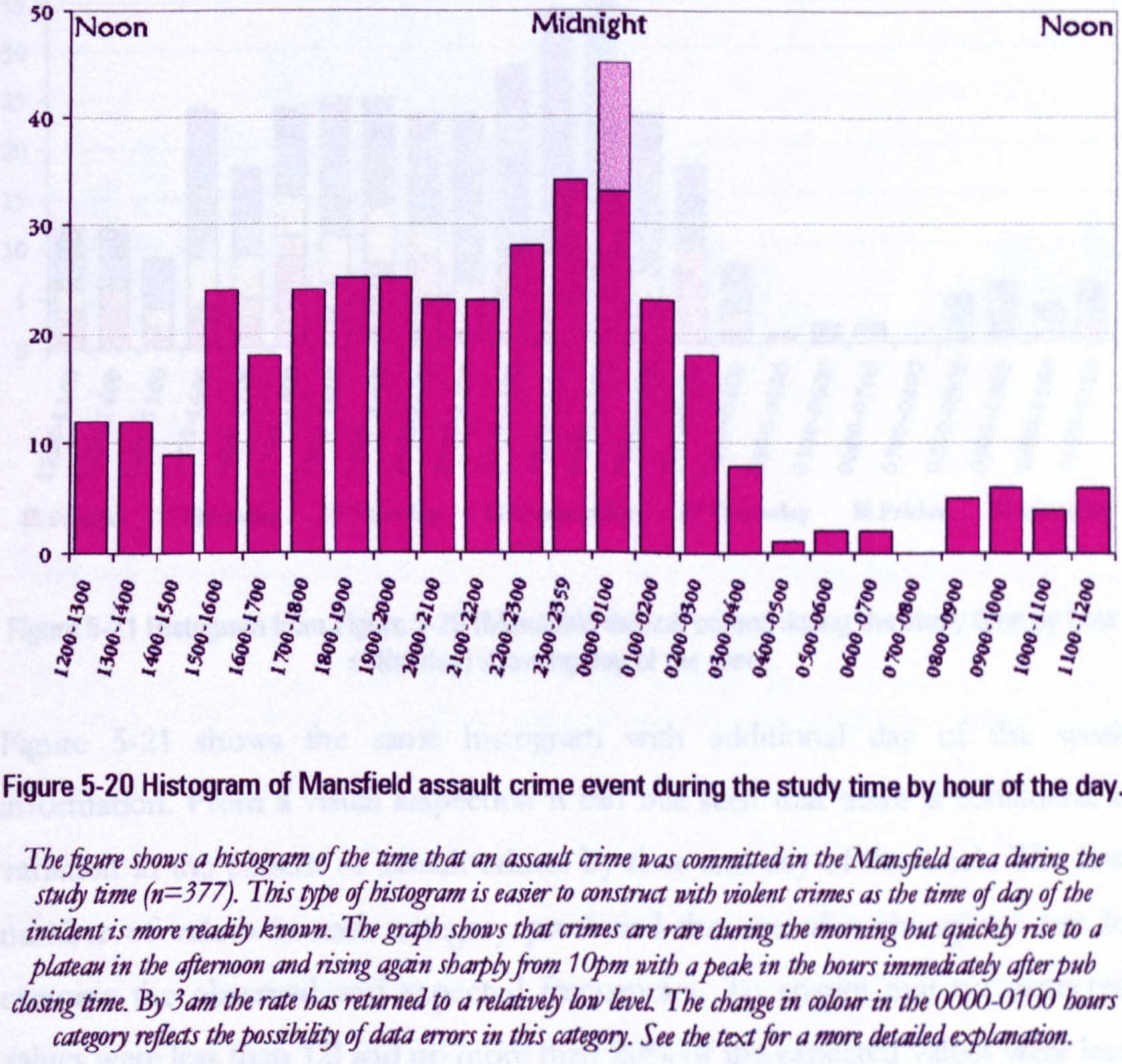


Figure 5-20 Histogram of Mansfield assault crime event during the study time by hour of the day.

The figure shows a histogram of the time that an assault crime was committed in the Mansfield area during the study time (n=377). This type of histogram is easier to construct with violent crimes as the time of day of the incident is more readily known. The graph shows that crimes are rare during the morning but quickly rise to a plateau in the afternoon and rising again sharply from 10pm with a peak in the hours immediately after pub closing time. By 3am the rate has returned to a relatively low level. The change in colour in the 0000-0100 hours category reflects the possibility of data errors in this category. See the text for a more detailed explanation.

One feature of the crime histogram is the high peak in crimes during the midnight to 0100 period. It should be noted that this may be a false peak caused by the crime recording methods of Nottinghamshire Constabulary. Of the 45 crimes during this time period 12 were recorded as having a time of 0000 hours. There are generally three causes for this. It is possible that a crime happened exactly at midnight though this is rare and officers often record the time of incident as 0001 for clarity. Secondly, cases of prolonged assault, such as a number of crimes that happen over the course of a day and are recorded on a single crime sheet, will not show any recorded time and will have 0000 entered in the time fields. Finally it is possible that the recording of 0000 hours is the result of data omission or data error. These possibilities are recognised in the graph with a lighter shade of purple for the top 12 values in the 0000 to 0100 hours category.

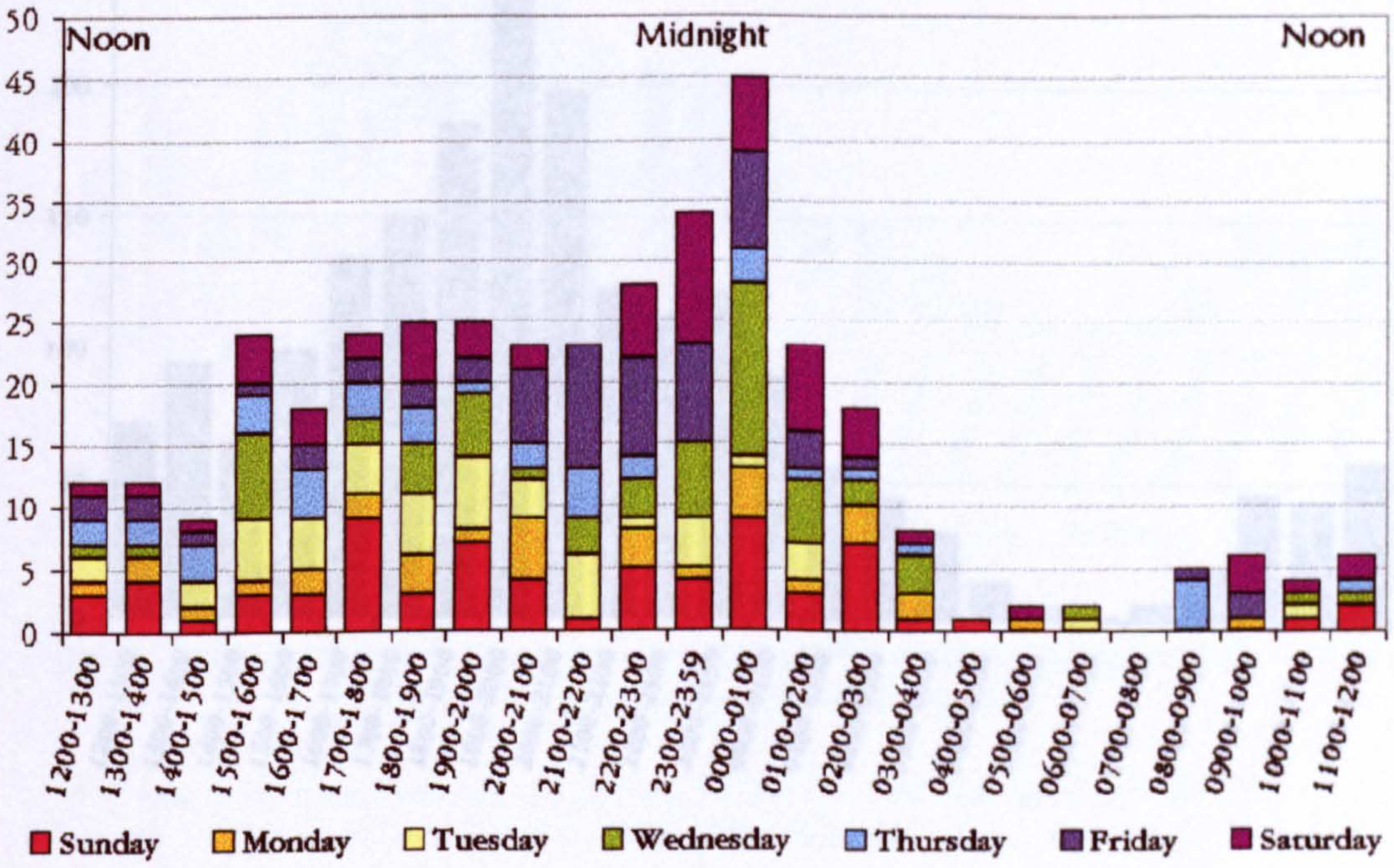


Figure 5-21 Histogram from Figure 5-20 (Mansfield assault crimes during the study time by hour of the day) showing day of the week.

Figure 5-21 shows the same histogram with additional day of the week information. From a visual inspection it can bee seen that there is considerable variation in the pattern of assault crimes by time and day of the week. The low number of values in each category precluded the use of a chi-square test to examine the observed and expected frequencies. To ensure that no expected values were less than 1.0 and no more than 20% of the expected values were less than 5 would have necessitated combining categories to such an extent that any test based on such broad categories would have had limited value.

The same horizontal scale as in the crime histogram is used in Figure 5-22 to show the histogram of incident calls, though with a different vertical scale. The potential for errors as in the 0000 hours problem described previously does not exist with incident data as the computer records automatically the time that an incident record was created. This is not to say that it is a record of when the incident took place, but reflects the time that the police were made aware of the incident. With modern communications and the widespread use of the 999 system this now tends to be an accurate (if slightly lagged) measure of the need for police assistance.

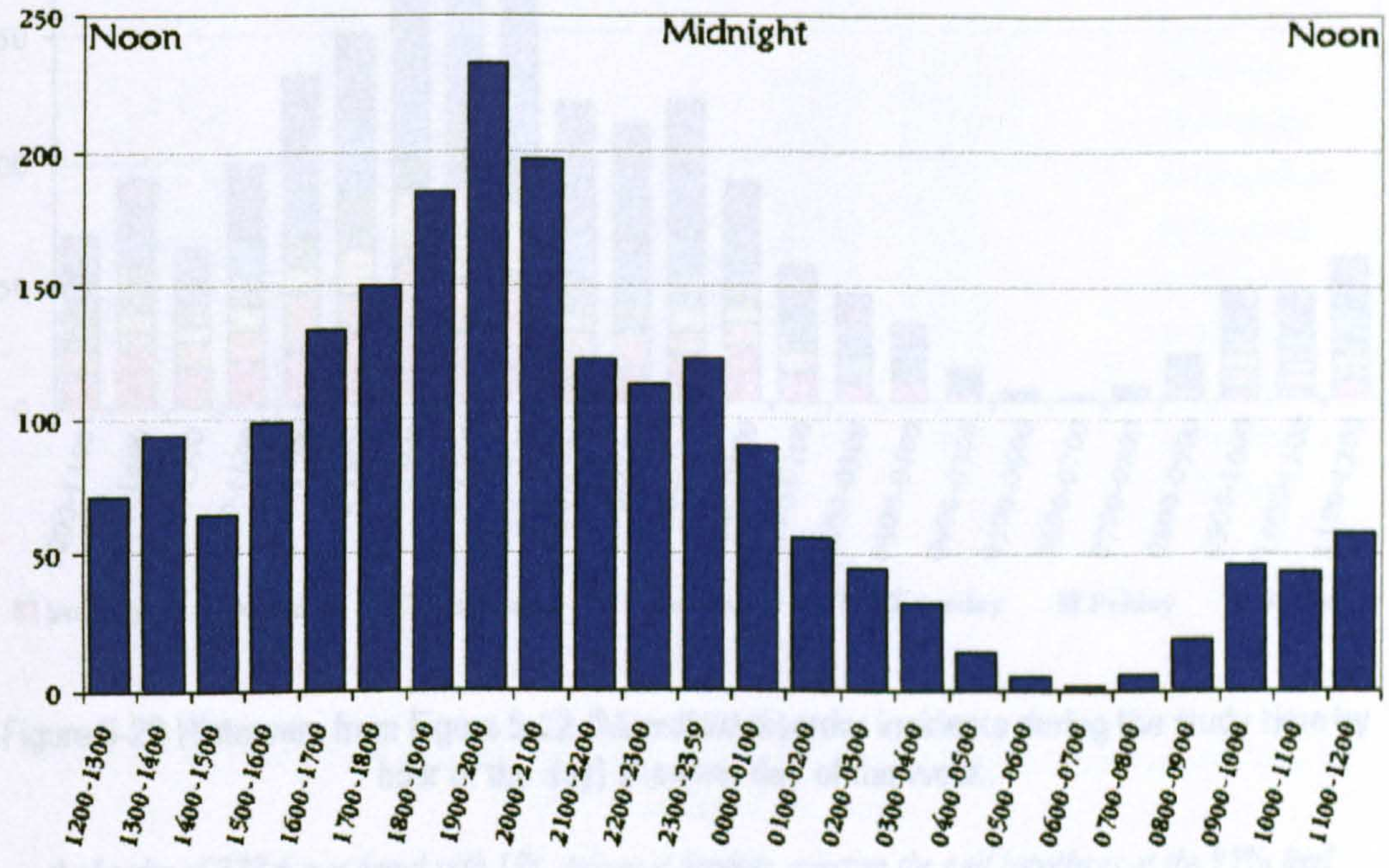


Figure 5-22 Histogram of Mansfield assault/disorder incident calls during the study time by hour of the day.

The graph shows the time of creation of incidents on the Nottinghamshire Constabulary command and control system (n=2000). This graph has a different vertical scale to the one shown in Figure 5-20 on page 128.

The incident data histogram is also plotted to consider the day of the week in Figure 5-23. The larger values in each category permits the application of a non-parametric chi square test when all incidents occurring between 0500 and 0800 are combined into one category. A χ^2 value of 222.6 with 126 degrees of freedom means that the null hypothesis (H_0: there is no significant difference in values by hour of day or day of week) is rejected at the 0.05 level (Neave, 1978). This shows there is a significant variation in the number of assault and disorder incidents in the Mansfield area, not only in the hours that the incidents occur, but there is also variation in the number of events by day of the week. Figure 5-23 shows variation in values both between days of the week and hour of the day. For example the figures for Tuesdays (yellow blocks in Figure 5-23) are noticeably more influential in the periods 1900-200 and 2000-2100 than earlier in the afternoon at 1200-1300.

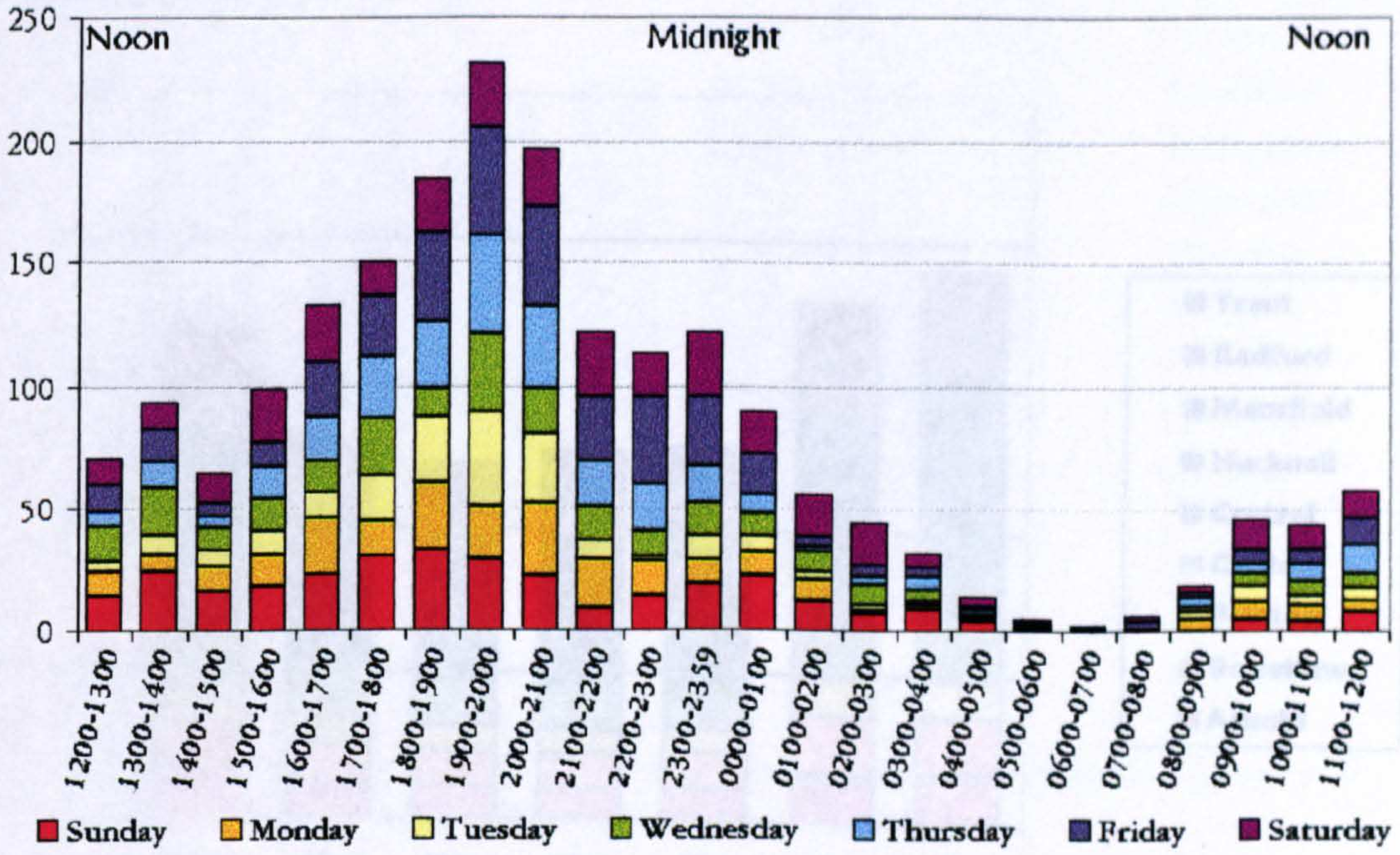


Figure 5-23 Histogram from Figure 5-22 (Mansfield disorder incidents during the study time by hour of the day) showing day of the week.

A χ^2 value of 222.6 was found with 126 degrees of freedom, rejecting the null hypothesis at the 95% level. Categories 0500-0800 were combined.

The graphs in Figure 5-20 and Figure 5-22 clearly demonstrate a different distribution of event times between the recording of incidents and the time of crimes. Calls for service (incidents) to deal with assault and disorder peak between the hours of 6pm and 9pm, while the peak for assault and disorder crimes is between 10pm and 1am. The distinctive peak found in the incidents is not reflected in the crime data. From this it would appear that any hypothesis to explain the number of assault crime reports as a function of the number of disorder incidents does not hold true in Mansfield for the period 1800 hours to 0100 hours, though it may be the case that the crimes lag the incidents by a number of hours. This possibility is examined later in this chapter.

5.6.1. Relationship with force-wide data

It is important to contrast the daily and hourly data extracted for Mansfield with that of Nottinghamshire Constabulary as a whole to confirm that Mansfield is a representative subset of the larger force area. This prevents the inadvertent extraction of useful intelligence about the picture in Mansfield which is not applicable to the other divisions.

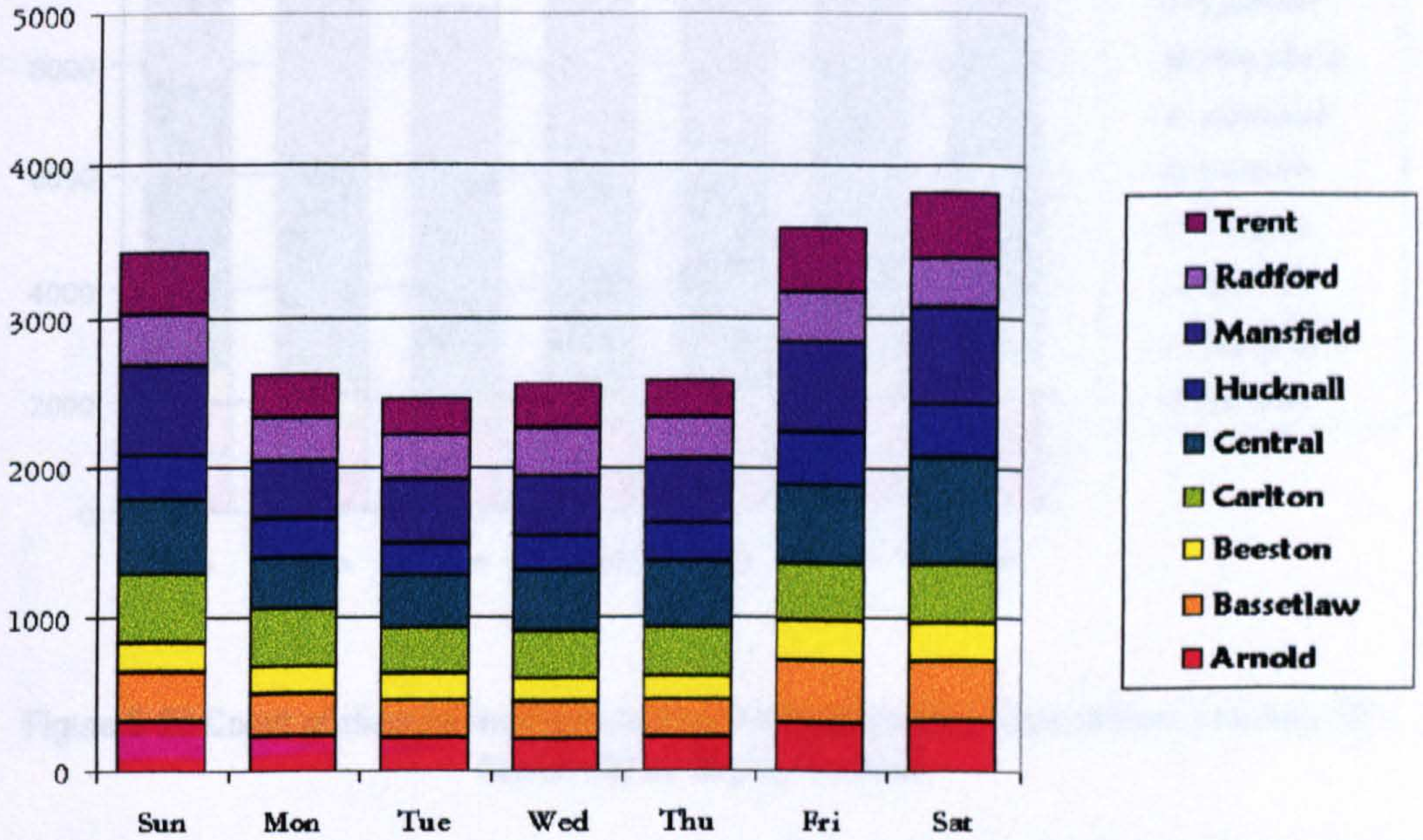


Figure 5-24 Count of violence crime reports across Nottinghamshire Constabulary (April 95 – April 97) by day of the week.

The TSA of Mansfield violence crime showed a noticeable weekend peak in the number of recorded crimes. This pattern is repeated in the data drawn from the whole of Nottinghamshire Constabulary (Figure 5-24 above). Saturday is the most populated class, though Sundays and Fridays are both noticeably larger than the midweek days.

The previous temporal analysis showed a definite weekend tendency in the incident data and this is replicated in the force-wide data (Figure 5-25). The whole weekend period (if Friday night is included) has more values than the midweek days and this corroborates the Mansfield analysis.

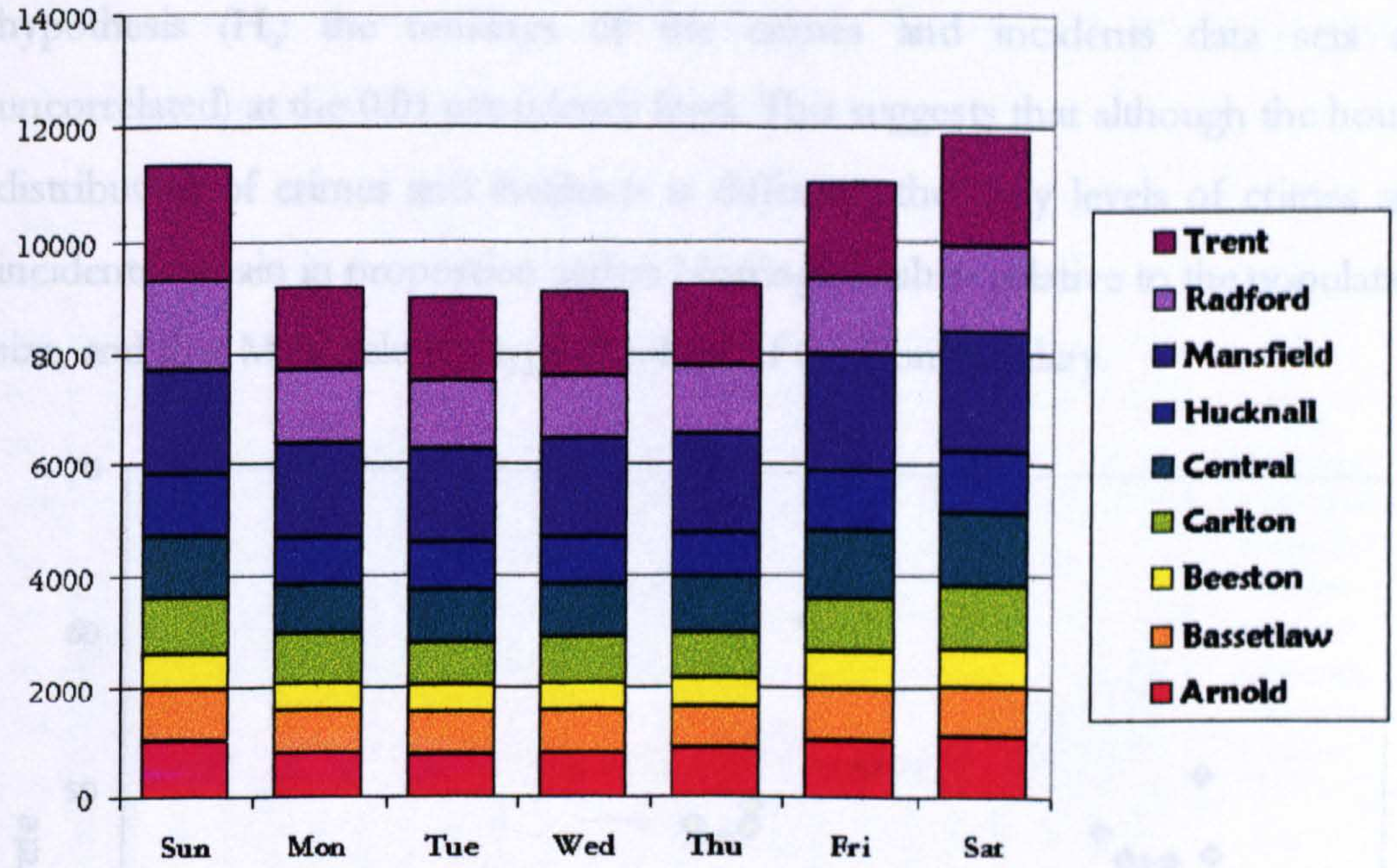


Figure 5-25 Count of disorder incidents across Nottinghamshire Constabulary (January 97 - March 98) by day of the week.

Note that data has been collected for a different period of time (January 97 – March 98) to that of the crime data (April 95 – April 97), and that the vertical scale is different from Figure 5-24.

Although at a different scale, the shape of the graph and the relative influence of each division in the overall total for a day in Figure 5-24 is similar to the distribution by day and division in Figure 5-25. This is reinforced by the scatterplot in Figure 5-26 that displays a high degree of clustering. This is to be expected as the two data sets are not independent and a certain number of incidents will generate a proportional number of crimes. This proportion of crimes to incidents is approximately the same for most divisions except Central and Radford which both display markedly different relationships between the variables. Central division displays a higher overall number of crimes and incidents with relation to the population. This is to be expected as the population of central Nottingham is low, the landuse of the area being predominantly retail and business. Radford division has a distinctly higher number of incidents compared to the rest of the county and this might be caused by a greater public awareness of crime and a higher number of calls to the police. The high correlation between the crime rates and number of incidents on each division on each day of the week in Figure 5-26 is confirmed by a Spearman Rho correlation coefficient of 0.74 (n=63) which rejects the null hypothesis (H_0: the rankings of the crimes and incidents data sets are uncorrelated) at the 0.01 confidence level. This suggests that although the hourly distribution of crimes and incidents is different, the daily levels of crimes and incidents remain in proportion across Nottinghamshire relative to the population size, and that Mansfield is a typical subset of the Constabulary.

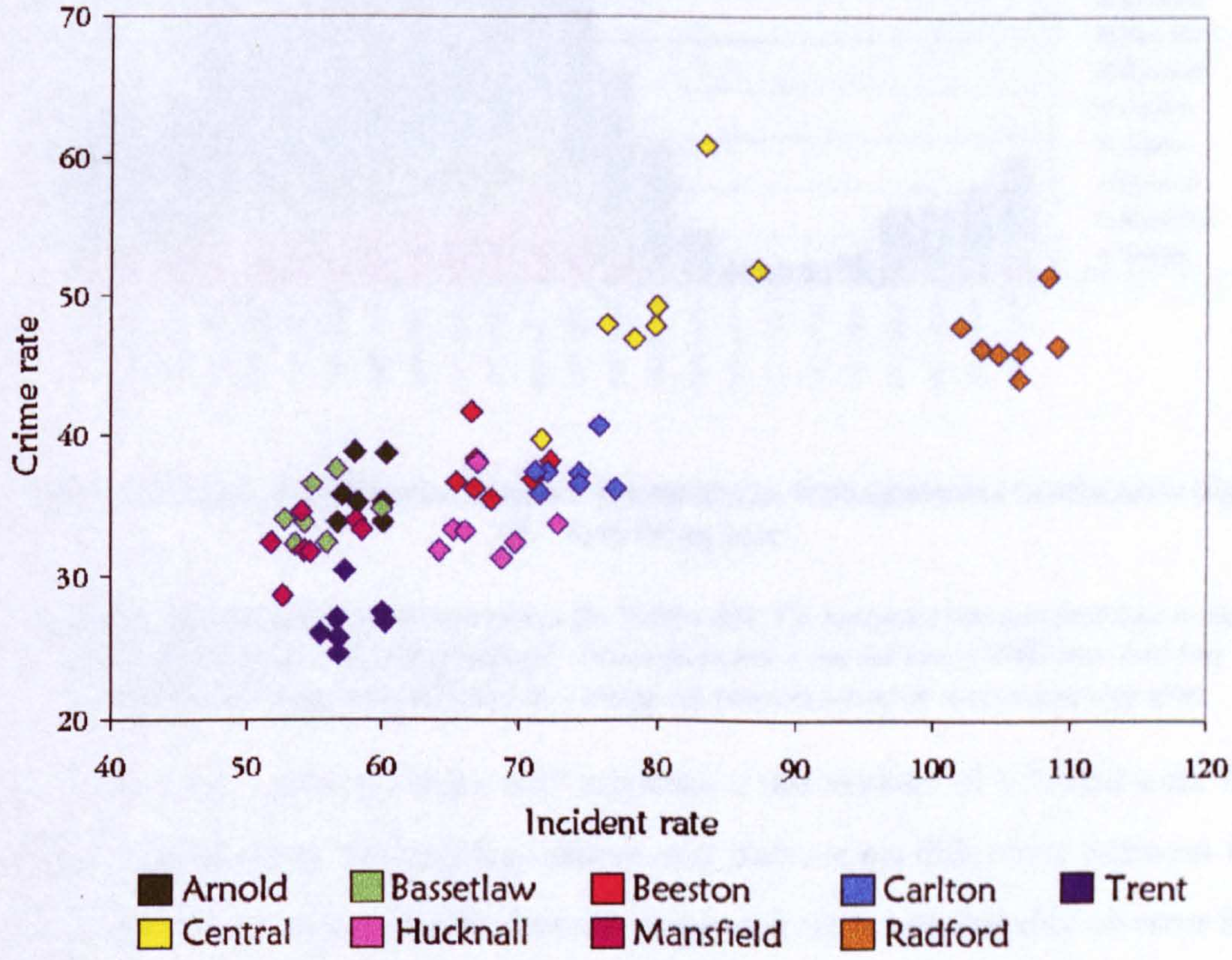


Figure 5-26 Scatterplot of the number of incidents and crimes per 1,000 residents on each day of the week on Nottinghamshire Constabulary divisions aggregated over January to March 1998.

Each point represents the number of incidents and crimes recorded for a division on a certain day. The correlation is confirmed by a Spearman Rho test which rejects the null hypothesis at the 0.01 confidence level.

HOURLY RATES

The Mansfield study extracted a report of the assault crimes and disorder incidents and created two histograms of this data (Figure 5-20 on page 128, and Figure 5-22 on page 130). These two graphs showed that incidents peaked a number of hours before the crime reports reached a peak, around pub and club closing time. The force-wide data extracted for all assault and disorder crimes and incidents confirmed the Mansfield pattern. Figure 5-27 shows the violent crimes histogram for the Constabulary area.

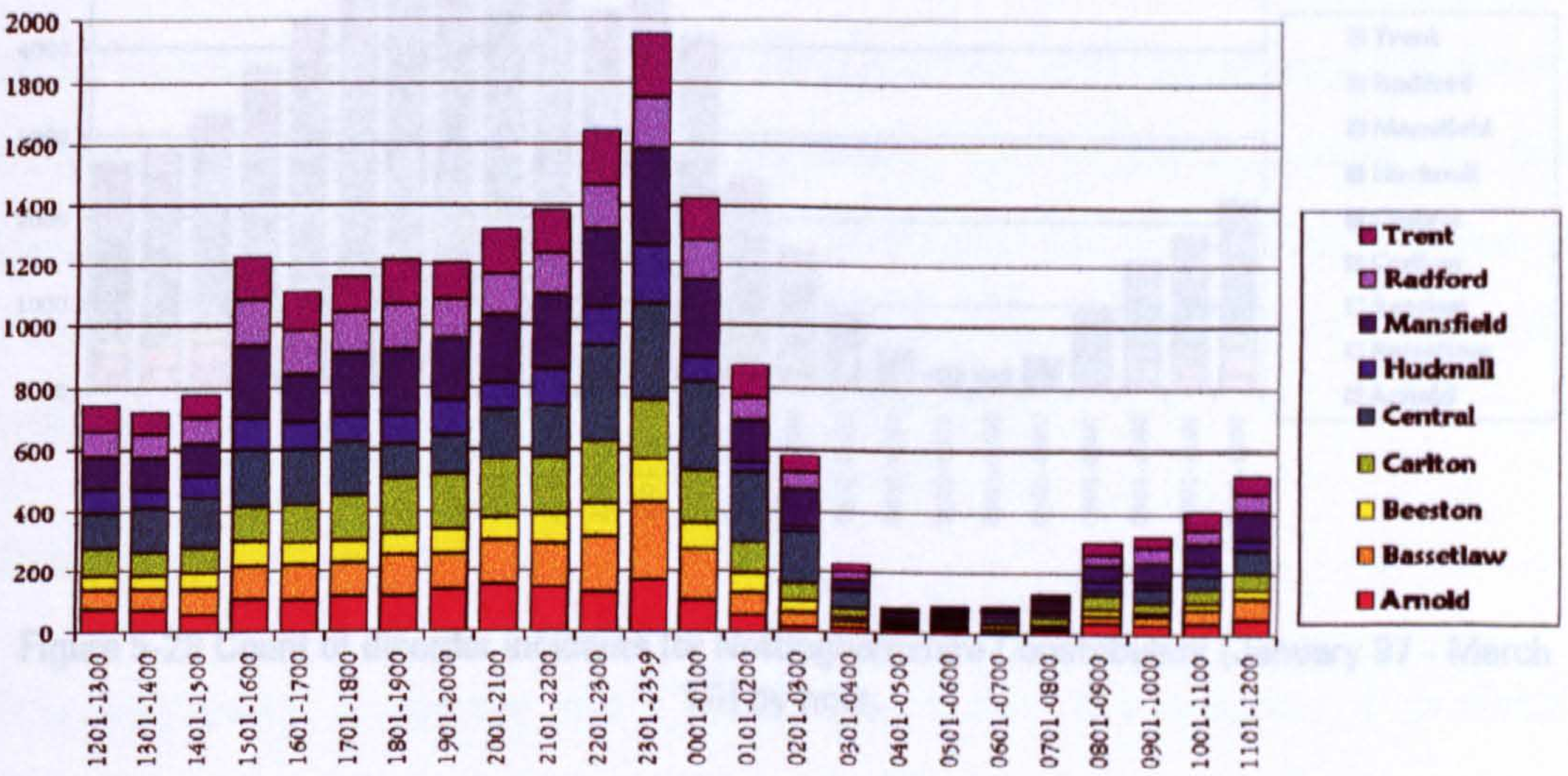


Figure 5-27 Count of violence and assault crime reports for Nottinghamshire Constabulary (April 95 – April 97) by hour.

The time of offence shown in the time recorded in the TOtime field. The horizontal scale runs from noon to noon with the highest peaks either side of midnight. Crime records with a recorded time of 0000 hours have been excluded due to the high probability that these timings are caused by incomplete data or data entry errors.

A χ^2 test of the data in Figure 5-27 returned a test statistic of 679 and with 184 degrees of freedom, the null hypothesis that there is no difference between the hourly levels of crime on each division is rejected with a probability of error less than 0.01. There is a significant difference between the divisions in the hourly rates of assault and other violent crime. An example of this can be seen in Figure 5-27 where from 1500 hours to 0100 hours the level of crime in Radford stays approximately the same, while over the same period the level of crime in Central and Bassetlaw divisions fluctuates. Figure 5-27 shows a noticeable rise in the recorded crime rate after 3pm (1500 hours) which remains high all afternoon eventually rising above the 1400 crimes rate for the three hour peak of 10pm (2200 hours) to 1am (0100 hours). This is a similar pattern to that seen in the Mansfield recorded crime data in Figure 5-20 on page 128.

The incident disorder data for Mansfield (Figure 5-22 on page 130) showed a peak much earlier in the evening between 6pm (1800 hours) and 9pm (2100 hours). This peak and the general pattern is also replicated in the force-wide data in Figure 5-28.

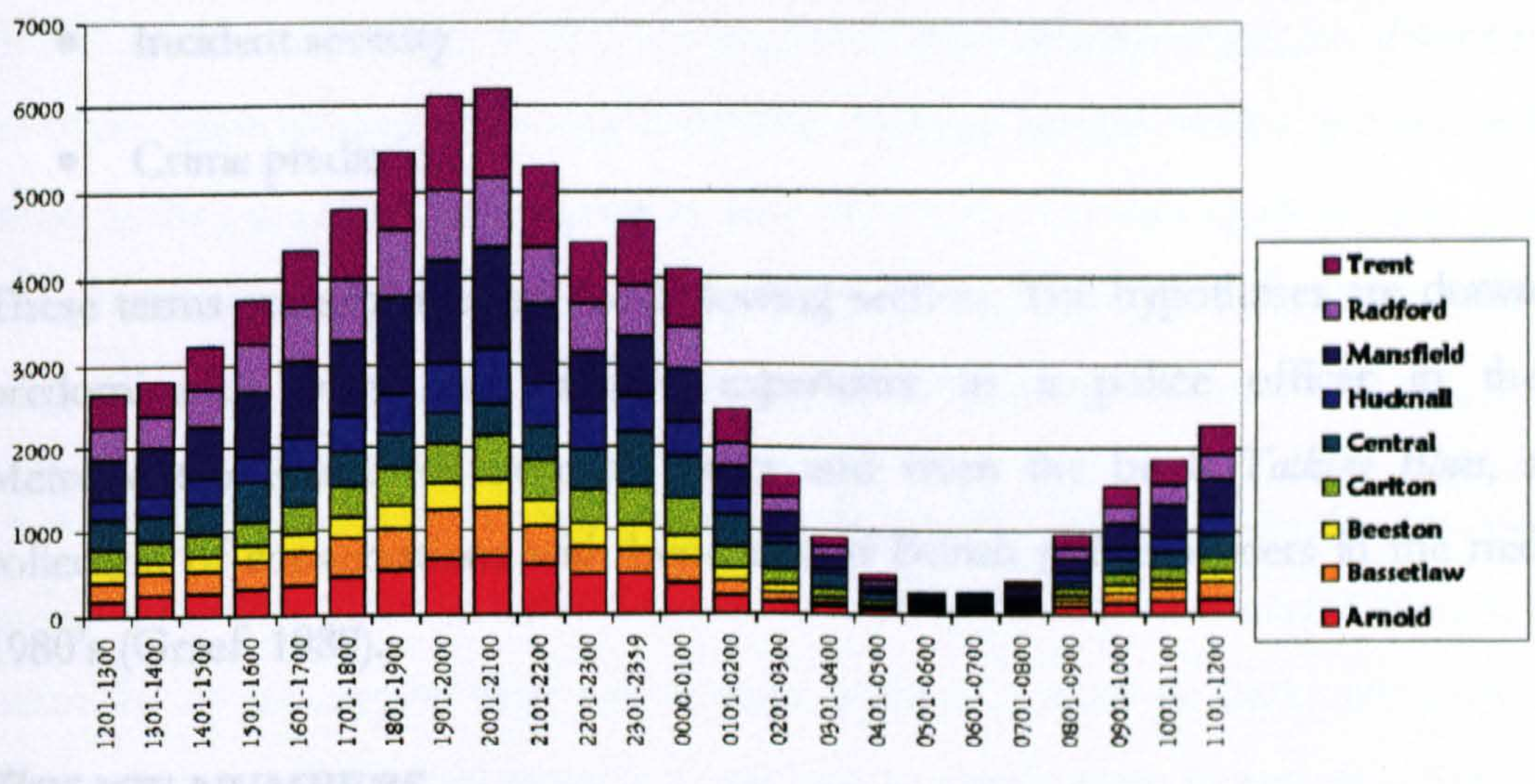


Figure 5-28 Count of disorder incidents for Nottinghamshire Constabulary (January 97 - March 98) by hour.

A χ^2 test of the data in Figure 5-28 returned a test statistic of 1439 with 184 degrees of freedom. The null hypothesis that there is no difference between the hourly levels of crime on each division is rejected with a probability of error less than 0.01. There is significant difference between the divisions in the hourly rates of disorder incidents.

The graphs and χ^2 tests in this section have shown that the variations in the Mansfield data are not isolated cases, and that the temporal variation in disorder incidents and crimes of violence exists force-wide across Nottinghamshire.

5.6.2. Explaining the daily temporal results

The previous sections have highlighted differences in the temporal distribution of assault and disorder crimes and incidents within Mansfield division and force wide. It is possible to hypothesise a number of reasons for this difference in distribution. A number of possible reasons include:

- Police numbers
- Policing style
- Delayed response
- Public tolerance

- Incident severity
- Crime prediction

These terms are explained in the following section. The hypotheses are drawn predominantly from the author's experience as a police officer in the Metropolitan Police for over ten years and from the book *Talking Blues*, a collection of conversations with hundreds of British police officers in the mid 1980's (Graef, 1989).

POLICE NUMBERS

One hypothesis to explain the distributions is that the number of available officers changes during the 24 hour period. This is normal with the changes in shift pattern as some shifts (commonly called 'reliefs' in police parlance) will have officers who are sick, on courses or annual leave. The number of remaining officers available to police the streets is constantly fluctuating though stations aim to provide a minimum number. The number of available officers will also change during the course of a shift as officers become unavailable while arresting criminals or being sent away from the division on errands. The distributions seen could be explained by the greater availability of officers after 2200 hours to patrol actively and generate arrests for disorder incidents. These extra arrests will appear directly in the crime data without necessarily appearing in the incident total as calls generated by the public.

POLICING STYLE

The style of policing employed in the local area may also be a cause of the distribution. An arrest for assault can cause an officer to be kept at the station for many extra hours beyond their normal shift. Many shift patterns in different forces end at 10pm when the night shift come on duty. Although there is a peak of calls for service just before 10pm (Figure 5-22) officers may be unwilling to make arrests close to the time they had hoped to go home and they may be exercising their discretion to find alternative remedies in situations where they would normally resort to a crime report and possible incarceration.

If the relief are short of available manpower then officers may also decide to exercise extra discretion to enable them to remain 'on the streets' for a longer time as they may feel that otherwise they will leave their colleagues to deal with the incidents in the area for the rest of the shift alone. Again in these circumstances they may try to find an alternative recourse in place of arrest.

A third possibility in the policing style category might be caused by the desire in officers to seek more 'action'. Amongst more experienced officers a mundane arrest holds less appeal than for younger officers in their probationary period. They may decide to avoid arrest in a situation to enable them to remain active on the streets. This would enable an active search for a more interesting and 'worthwhile' arrest. This can be noticed in the case of officers known locally as a 'thief-taker'. There is considerable kudos in being known as a thief-taker though this can not be achieved by dealing with routine calls. Officers seeking a positive reputation will avoid mundane calls and making arrests for uninteresting work in order to concentrate on the more glamorous action.

DELAYED RESPONSE

When the station is receiving more calls than they can deal with, the response time to incidents starts to grow and officers will feel under more pressure to assist their colleagues. If they respond to a call for service and a suspect for an assault is named who lives nearby, they may decide to call on the suspect at a later quieter time. They may do this if they feel there is a good possibility of the suspect being present at that later time and an arrest being made. Knowing that they will be off the street for hours during a busy period can influence their decision-making. They might also decide not to call on the suspect but to pass the information on to the next relief to allow them to effect the arrest.

PUBLIC TOLERANCE

The differences in the distributions in Figure 5-20 and Figure 5-22 are most noticeable between early evening and late evening/early morning. Another cause of this difference might be the result of public tolerance. The importance of the public in generating information to the police has been stated earlier and the number of incidents has been described as a measure of the public need to seek an authoritarian remedy to an undesirable situation (Morgan and Newburn, 1997). It is possible that the peak in calls during the early evening is a reflection of the lower tolerance of the public to minor misdemeanours at the end of the day. If people have been at work and away from the area they may be unused to the disruptive behaviour of local youths, or unwilling to tolerate their behaviour. The lack of an increase in recorded crime may be a recognition that the police are more used to minor infringements of the law and they are less likely to record the incidents as crimes.

INCIDENT SEVERITY

Although there are fewer calls after 9pm there is an increase in recorded crime. It seems likely that although the calls to the police decrease they increase in severity. The influence of alcohol as a cause of delinquent acts has been documented elsewhere (South, 1997) and this may be a factor in the possible increase in incident severity, prompting a crime report. If reported incidents increase in severity then fewer calls will be needed to generate an increase in crime reports.

CRIME PREDICTION

A final hypothesis is based on the notion that more trivial incidents during an evening are a precursor to more serious incidents later in the evening. This hypothesis opens up the possibility that the location and number of incidents in the early hours of the evening can give an indication as to the location of more serious assaults at a later time. This would require a correlation between the location of the incidents and crimes, with a time lag in the short term.

5.6.3. Taking the analysis further

The previous sections have generated a number of questions regarding the relationship between crime records and calls for service. The seasonality components combined with the correlation analysis have identified the weekend distribution of crimes and incidents as an interesting source of further work. The histograms of daily temporal patterns have also raised a number of possible hypotheses to explain the differences in distribution. While the possibility of using incidents to predict crime in the short term is interesting, it can only be examined to a moderate extent with the current data sources as they only overlap for a three month period (January to March 1997) and are not truly independent. However an exploratory analysis is possible and the next section of this chapter extends the analysis to include the geography of the crime and incident sites in an attempt to further explain this complex relationship.

5.7. FURTHER ANALYSIS

The following section aims to address three of the questions raised by the previous analyses. These questions arise from the different temporal distribution of incidents and crimes, and the importance of the weekend data distribution in analysing violence and disorder in Mansfield. The questions are:

1. Is there a discernible geographical pattern to the distribution of incidents and crimes across Mansfield?

2. Does the distribution in the first question reveal a different pattern over the weekend?

3. Do incidents that occur in the early evening time period share a general geographical location with crimes that occur later in the same evening?

5.7.1. The geographical distribution of incidents and crimes

The original data was re-examined and weekend incidents were identified. The original 377 crimes and 2000 incidents were mapped to a one kilometre grid superimposed over the Mansfield police divisional area. This grid is shown overleaf in Figure 5-29 with the main urban centres of the region.

The modifiable areal unit problem has been discussed in the previous chapter (4: Aoristic crime analysis) and the limitations of the MAUP are well-known. Larger grid squares would aggregate the data further, resulting in too general a study, and a finer grid mesh would result in most of the squares containing so few points that any results would lack any analytical value.

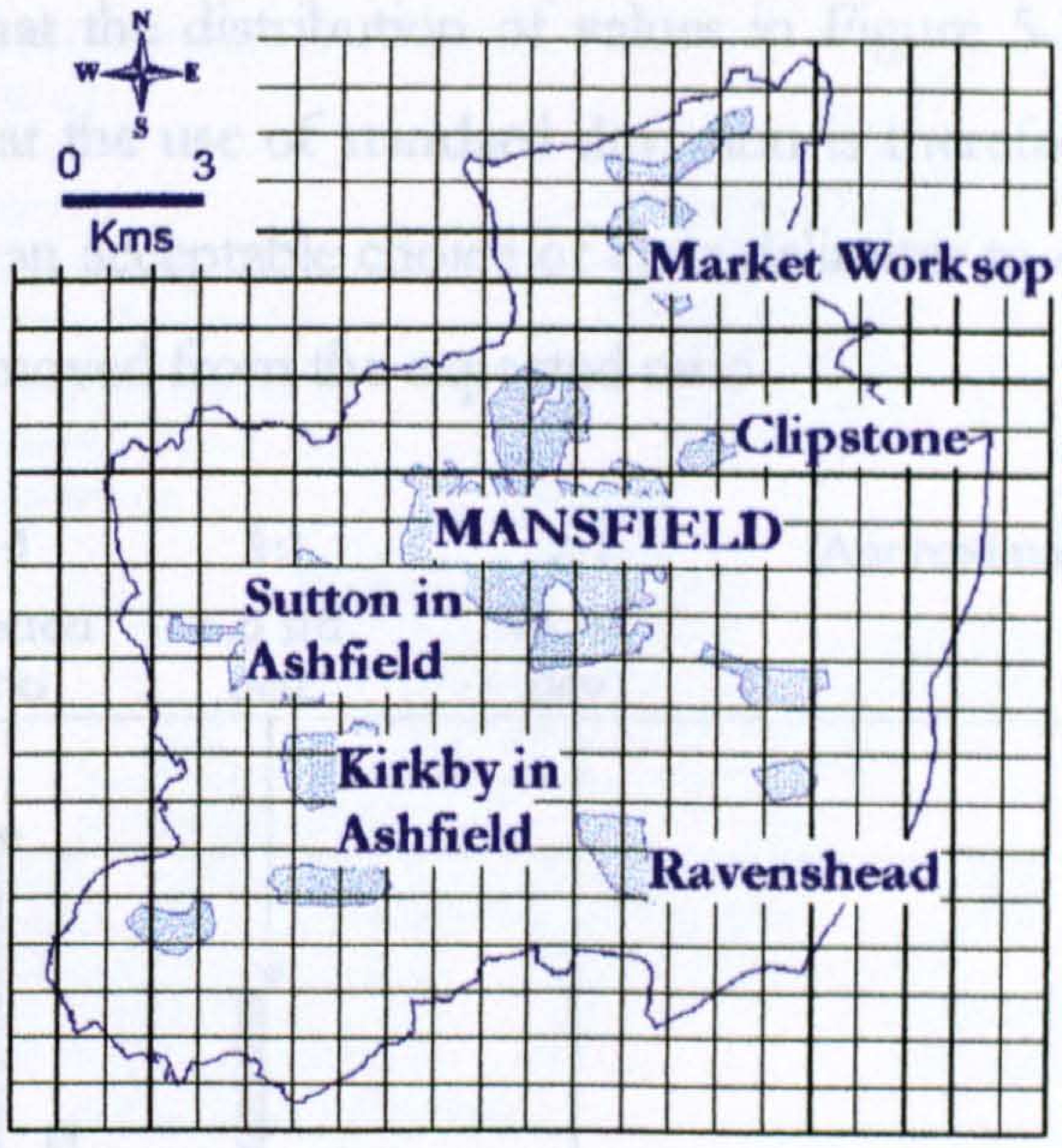


Figure 5-29 Mansfield urban areas overlaid with a one kilometre grid.

The graphs in Figure 5-20 and Figure 5-22 show that there is a different temporal distribution in these series and it may be that there is a different spatial pattern. There are 377 recorded crimes and 2000 recorded incidents uniformly spread over the study area. 56 cells contain only incidents and four cells contain only crimes leaving 73 grid cells which contain both crimes (n_{crimes}=372) and incidents ($n_{incidents}$=1782). If these cells are examined we find that the expected incident to crime ratio within these cells becomes approximately 5:1. This removes the distortion created by cells which only contain one type of event. A histogram of the 73 cells which contain incidents and crimes is shown in Figure 5-30. A standard deviation from the expected ratio of approximately 5:1 (0.209) was calculated as 0.298 (nearly 3:1). Standard deviations from the expected ratio are shown in Figure 5-30 at 0.5 deviations. Ratio values higher than the expected ratio indicate more crimes than expected and are coloured purple beyond 0.5 deviations. Ratio values less that the expected value indicate a greater proportion of incidents than expected, and these are coloured blue when less than 0.5 deviations from the expected ratio. The majority of the cells are within half a standard deviation of the expected ratio (coloured grey). The large outlier value at 1.0 is caused by 6 grid cells which contained one crime and one incident (a ratio of 1.0).

It is appreciated that the distribution of values in Figure 5-30 is not a normal distribution and that the use of standard deviation is therefore debatable, but it was felt that it was an acceptable choice of class delimiter to demonstrate simply the values more removed from the expected ratio.

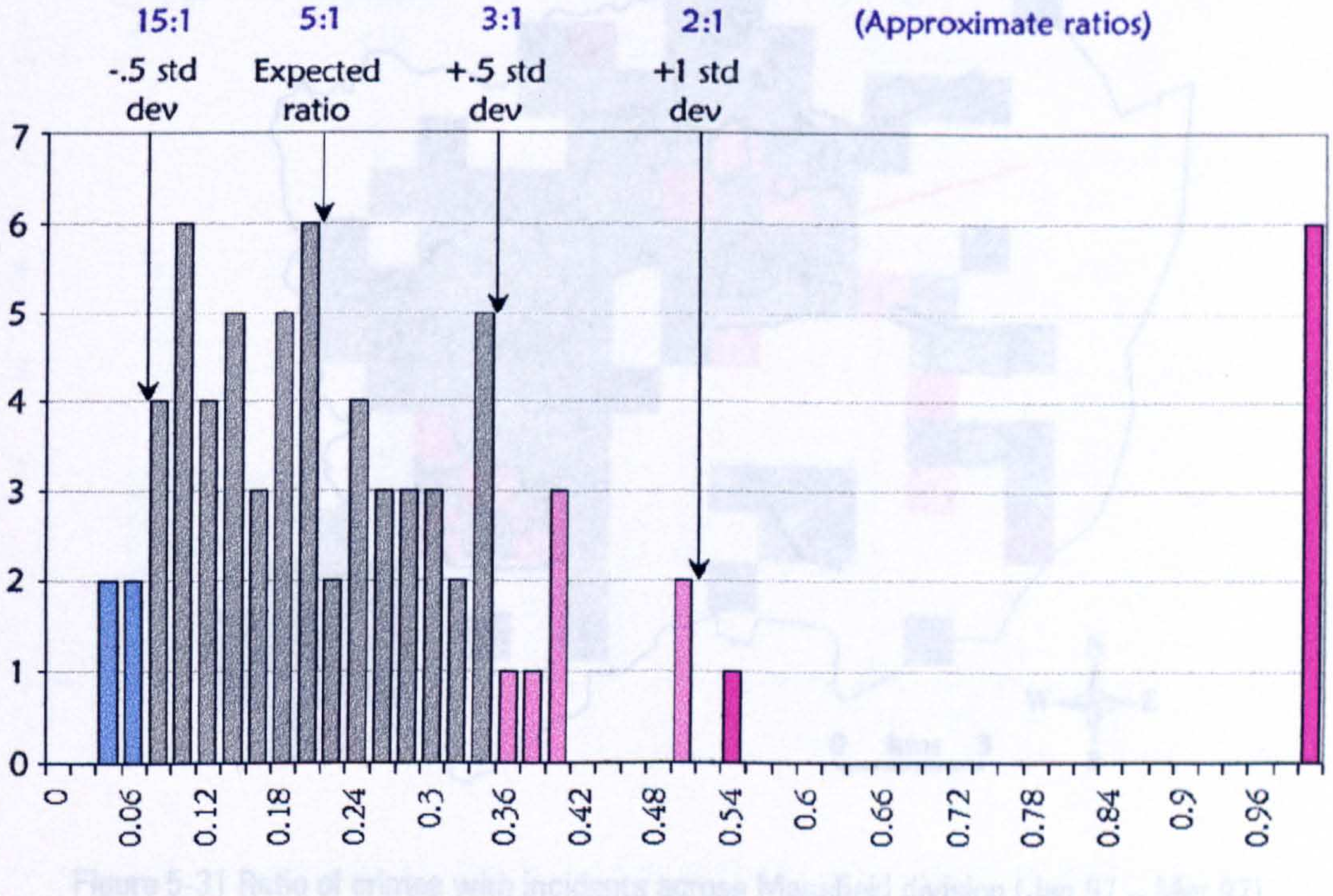


Figure 5-30 Histogram of Mansfield incidents to crimes ratio of those grid cells contain both data types.

Values representing cells within half a standard deviation are coloured grey in the graph, and in the maps to follow. A greater than expected ratio of crimes to incidents is indicated by the purple classes, and blue is used to represent classes with a greater than expected ratio of incidents to crimes Actual values are 0.209 (expected ratio) and 0.298 (standard deviation). The outlier at 1.0 is caused by six cells with one crime and one incident.

The individual cells from Figure 5-30 are plotted over the page in Figure 5-31 along with the cells which contained only crimes or only incidents. No distinct pattern emerges from the ratio of crimes and incidents in Figure 5-31, though it is clear that much variation exists. The grey cells close to the expected ratio are concentrated in or near the urban areas. A large number of cells (56) contain incidents without any corresponding crime reports appear to be on the periphery of the urban regions. The highest of these contains 23 incident reports but no recorded violent crime, and is indicated on the diagram with a red arrow. It can be found at the southern end of the Mansfield urban area.

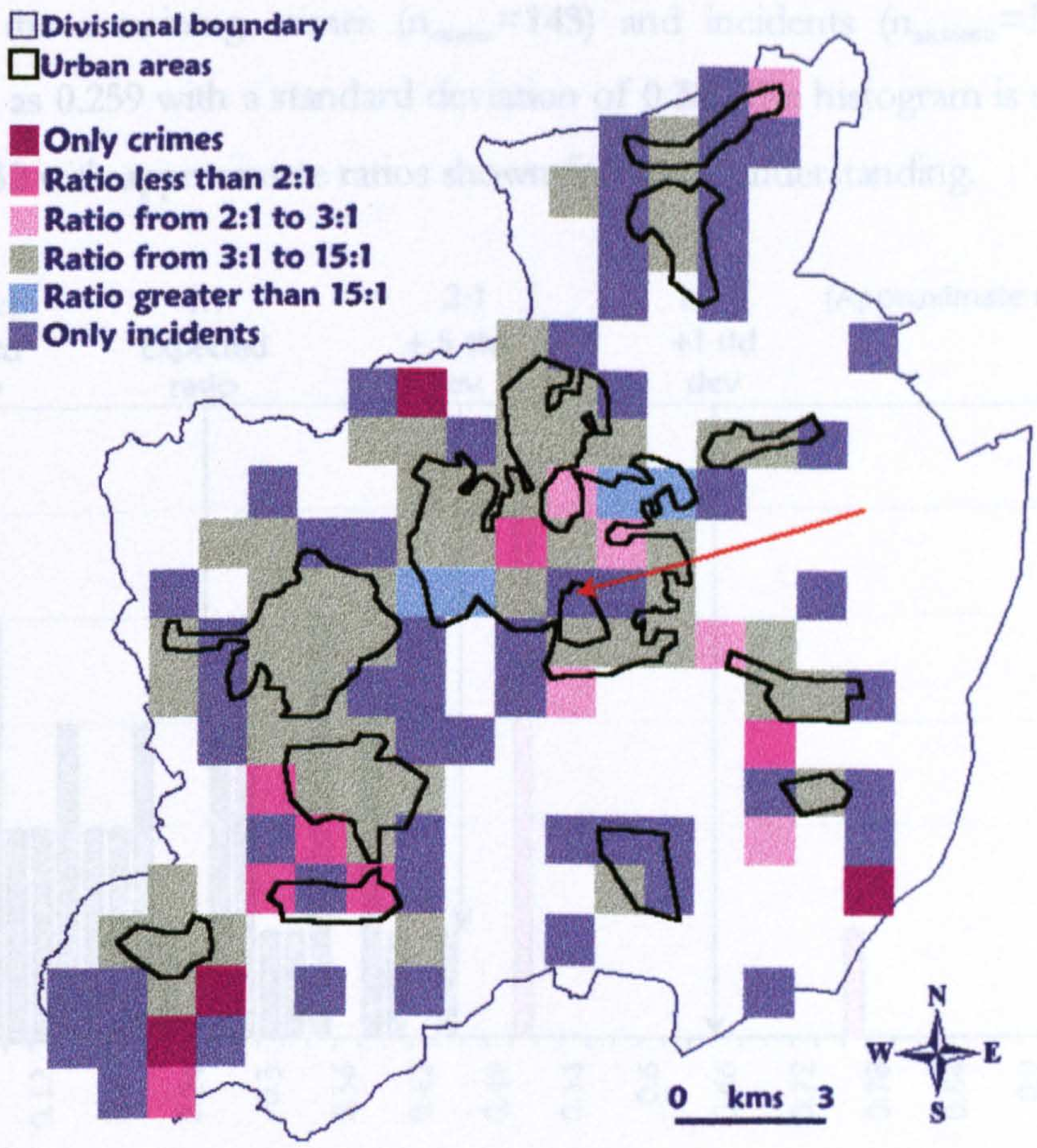


Figure 5-31 Ratio of crimes with incidents across Mansfield division (Jan 97 – Mar 97).

The figure shows all one kilometre square grid cells which had either a crime or an incident in the cell during the study period. Once cells with only one type of event were excluded, the expected ratio of the remaining cells was 0.209. A standard deviation of the cells containing both types of event was calculated as 0.298. The red arrow indicates the cell with the highest number of incidents (23) and no crime reports The same colour scale is used both here and in the histogram in Figure 5-30.

WEEKEND DISTRIBUTION

To examine the two series for weekend geographical variation, extraction of the weekend crime and incident records was necessary. As stated earlier, the prevalence of evening crimes and incidents meant extending the description of the weekend to include a Friday night. All events were selected which occurred between 6pm on a Friday night and 6pm on a Sunday night, allowing a full 48 hour period of offences and incidents. With only 148 crimes and 746 incidents occurring during this time, the analysis is more restricted than for the previous histogram and map of the distribution across the whole week. Once the cells containing only crimes (n_{cells}=5) and only incidents (n_{cells}=193) were plotted, the ratio for the remaining crimes (n_{crimes}=143) and incidents ($n_{incidents}$=553) was calculated as 0.259 with a standard deviation of 0.38. The histogram is shown in Figure 5-32 with approximate ratios shown for easier understanding.

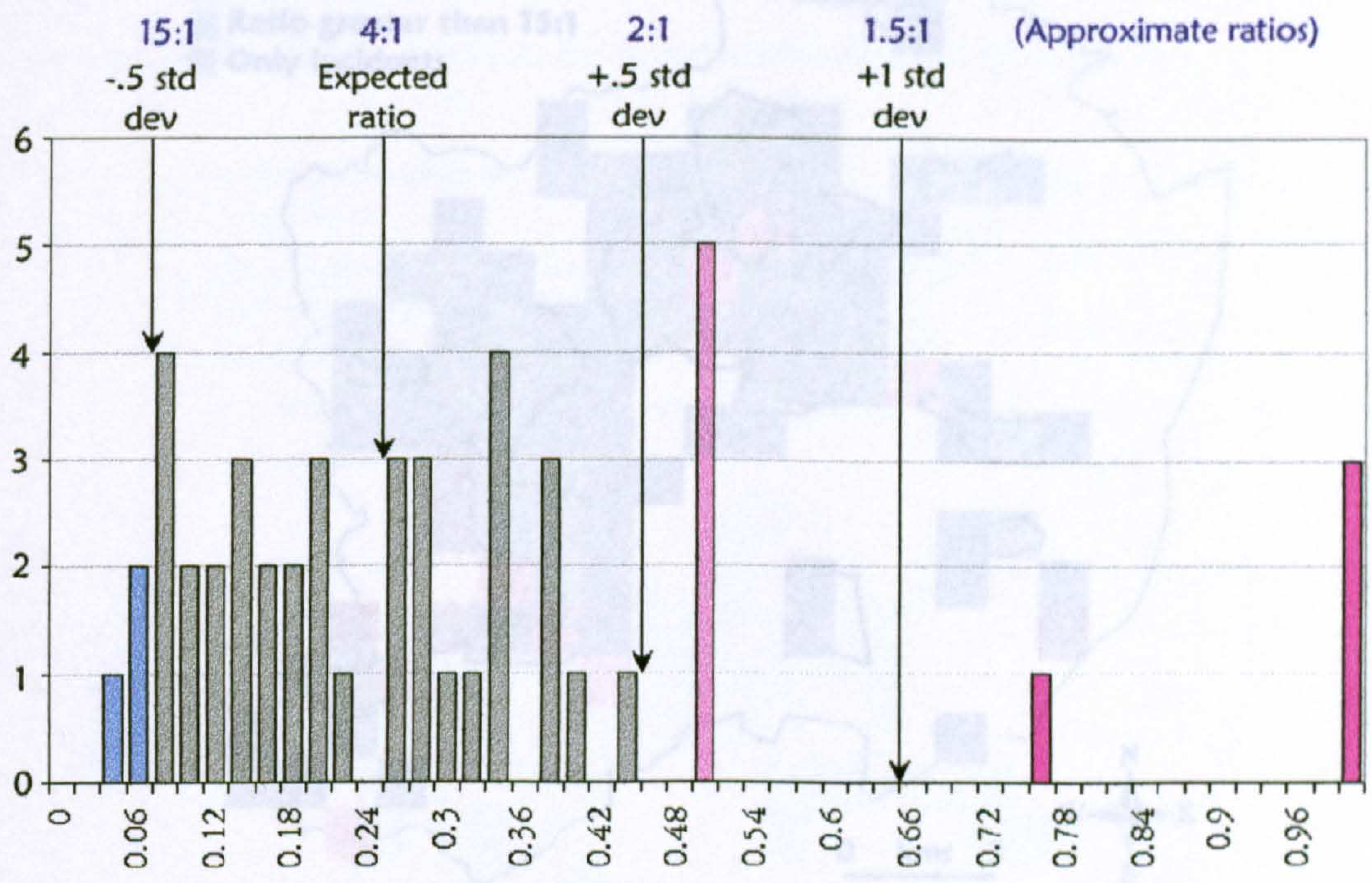


Figure 5-32 Histogram of weekend (Friday 6pm to Sunday 6pm) Mansfield incidents to crimes ratio of those grid cells contain both data types.

Values representing cells within half a standard deviation are coloured grey in the graph, and in the maps to follow. A greater than expected ratio of crimes to incidents is indicated by the purple classes, and blue is used to represent classes with a greater than expected ratio of incidents to crimes Actual values are 0.259 (expected ratio) and 0.382 (standard deviation). The outlier at 1.0 is caused by six cells with one crime and one incident.

The geographical distribution of all weekend grid cells containing either a crime or incident value is displayed in Figure 5-33, where a grid with the same georeferencing and dimensions as Figure 5-31 is employed. There is no one cell with a distinctly high incident number and so this image does not have a red arrow as in Figure 5-31.

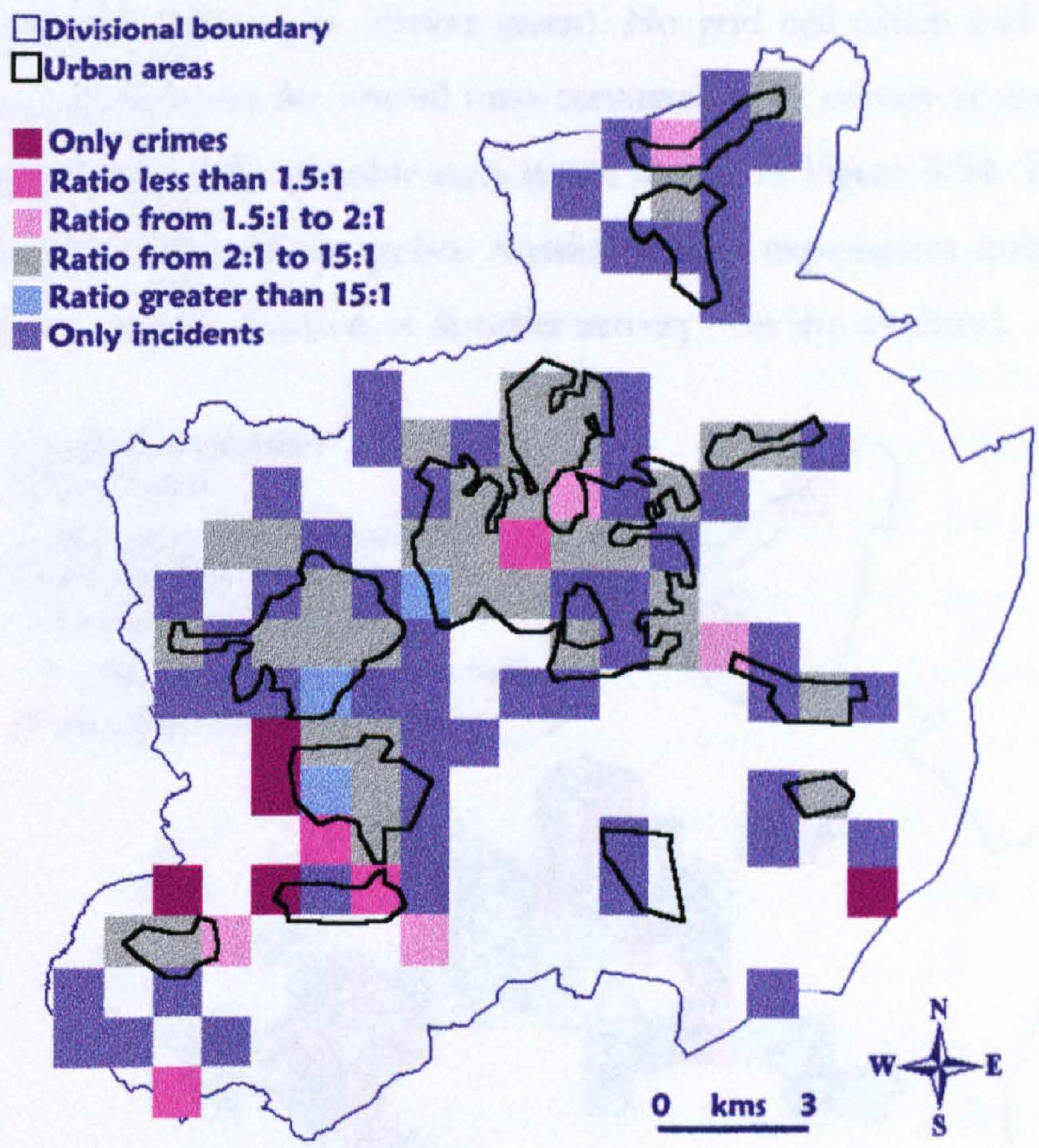


Figure 5-33 Ratio of weekend crimes with incidents across Mansfield division.

The figure shows all one kilometre square grid cells which had either a crime (n=148) or an incident (n=746) in the cell during the study period. Once cells with only one type of event were excluded, the expected ratio of the remaining cells was 0.259. A standard deviation of the cells containing both types of event was calculated as 0.362. A weekend event is defined here as occurring between 1800 hours on a Friday evening to 1800 hours on a Sunday evening.

COMPARISON OF WEEKEND AND OVERALL RATIOS

Figure 5-34 on page 147 shows the change from the overall map to the weekend-only map. Only those cells which had a ratio in the overall image (and therefore contained both crimes and incidents) are shown. To show the change in a simple coherent manner, there are four states indicated. If the ratio of crimes to incidents increases then the cell is shown in red. This would be an undesirable state as a crime is generally considered to be more serious than an incident. The ratio can remain the same, though this is the case for only two cells (shown in yellow). The two more desirable states are shown in green, where there has been a decrease in the crimes to incident ratio, or where the number of crimes within the cell has reduced to zero (darker green). No grid cell which had contained crimes and incidents in the overall map contained only crimes or no events at the weekend and so this possible state is not shown in Figure 5-34. This means that there is no area of the police division which experiences only weekday crimes or a complete cessation of disorder activity over the weekend.

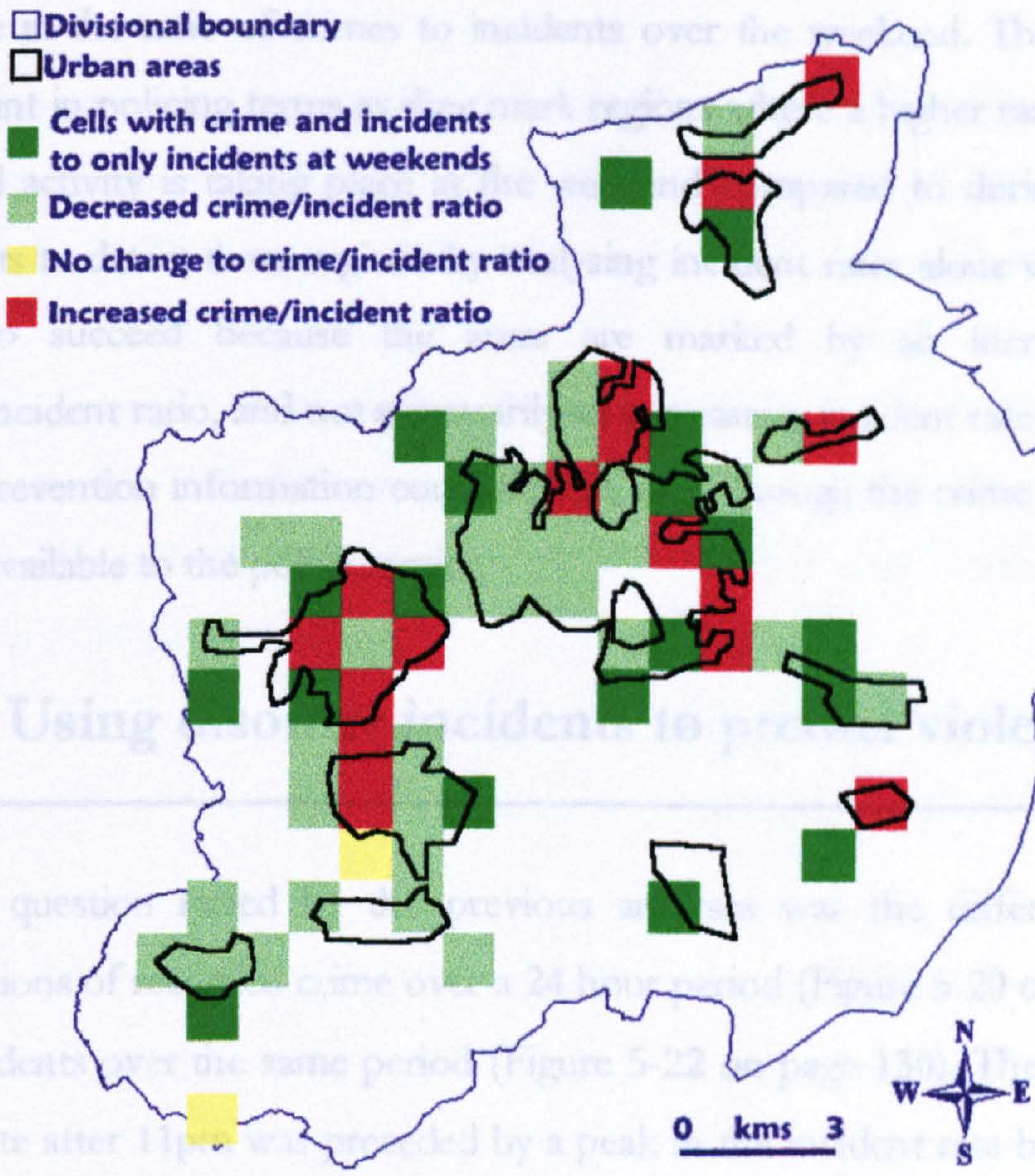


Figure 5-34 Comparison of cell change of state from all week to only weekends.

Only cells which contained crimes and incidents in the overall weekly map are depicted.

From Figure 5-34 it can be seen that there is an overall improvement in the ratio of crimes to incidents over the weekend when compared to the whole week, though a number of areas do show a deterioration. The overall improvement in ratio should be contrasted with the evidence from the rest of the analysis in this chapter that points to increased actual numbers of incidents and crimes. This would suggest than in the majority of areas in Figure 5-34 there is an increase in the *daily* rate of incidents over the weekend that is not matched by a proportional increase in the number of crimes.

With only 148 crimes and 746 incidents in the weekend data sets, one must be cautious when interpreting the analysis, however it does raise questions about the distribution of weekend crime and incidents, and a further examination should be undertaken if data for a greater temporal period becomes available.

The red areas of the map in Figure 5-34 show grid cells where there was an increase in the ratio of crimes to incidents over the weekend. These areas are important in policing terms as they mark regions where a higher ratio of serious criminal activity is taking place at the weekend compared to during the week. Attempts to detect these regions by analysing incident rates alone would be less likely to succeed because the areas are marked by an increase in the crime/incident ratio, and not necessarily an increase in incident rates alone. Vital crime prevention information could be lost, even though the crime and incident data is available to the police service.

5.7.2. Using disorder incidents to predict violent crime

A final question raised by the previous analyses was the difference in the distributions of recorded crime over a 24 hour period (Figure 5-20 on page 128), and incidents over the same period (Figure 5-22 on page 130). The peak in the crime rate after 11pm was preceded by a peak in the incident rate between 7pm and 8pm. An analysis was undertaken to see if the incidents happening in the early evening were precursors of crimes near that location later in the night.

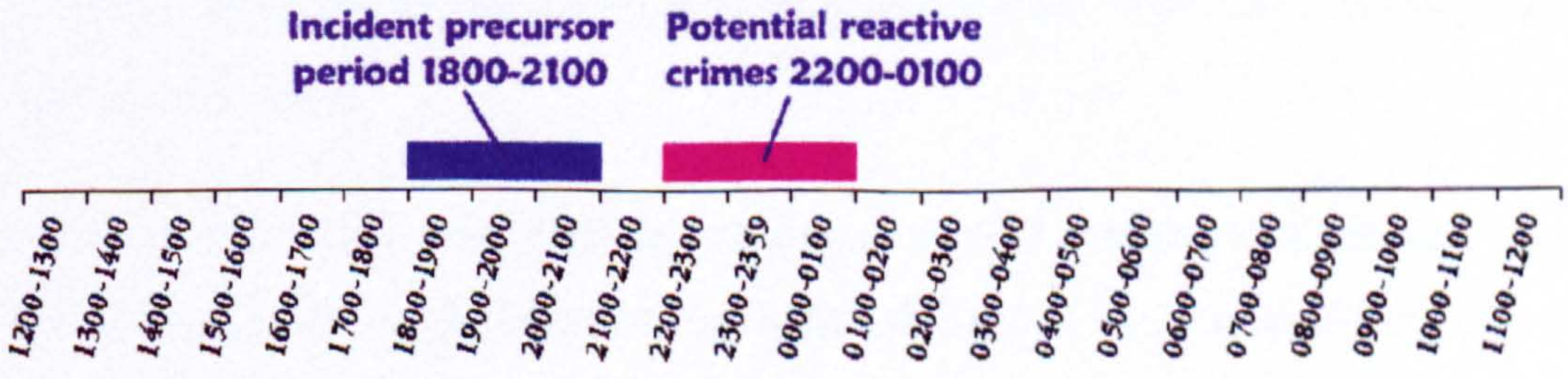


Figure 5-35 Time of searches for predictive incidents and reactive crimes.

The hours used in the predictive model were expanded to two three-hour periods. The search for precursor incidents was extended to the period from 6pm to 9pm (1800-2100 hours) and the search for reactive incidents ran from

10pm to 1am (2200 to 0100 hours) as in Figure 5-35. This more expansive search found 113 crimes and 618 incidents over a 90 day period. These time periods were chosen because they represented high values in the hourly distribution graphs (Figure 5-20 on page 128 and Figure 5-22 on page 130). Leaving a gap of an hour between the search times reduced the possibility of an incident and a crime being the same event.

Animation was used to visualise the results of this analysis. The process of animating results was discussed in depth in the previous chapter (4: Aoristic crime analysis). The program 'Dave's Targa Animator' was used to generate an FLI file called **CRIME PREDICTION ANIMATION.FLC**, which can be found on the accompanying CD-ROM. This format is playable with AAWIN.exe, also included on the CD-ROM. A collection of still images from the animation are shown in Figure 5-36. Note that the full animation is best viewed at a screen resolution of 800x600 or better.

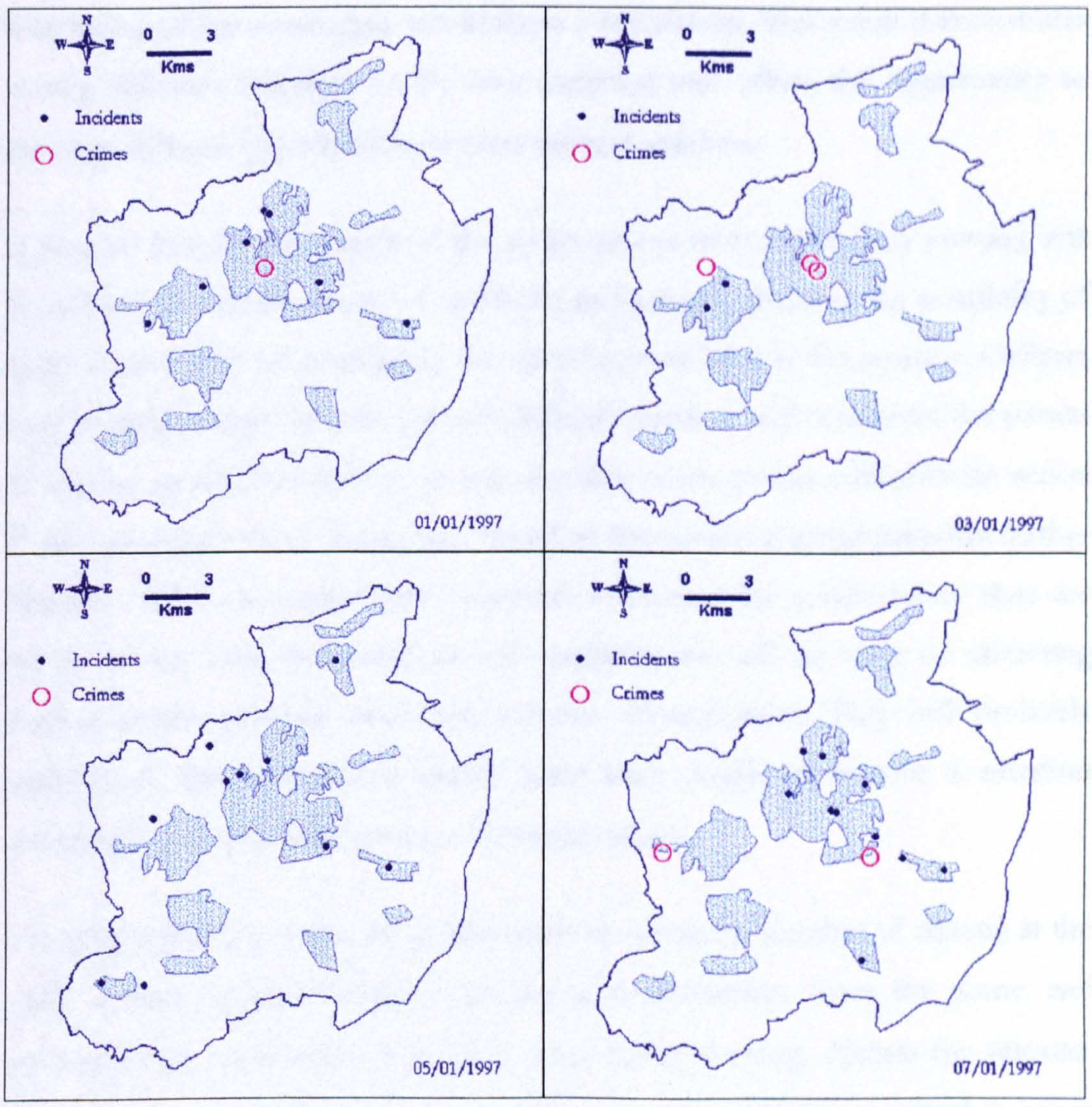


Figure 5-36 Collection of frames from Crime_prediction_animation.flc animation sequence.

Each of the animation frames shows the Mansfield police boundary, and within this the urban areas as shaded blue regions. The image date is shown in the bottom right of the frame. The potential precursor incidents are shown as solid blue dots while the purple circles indicate possible reactive crimes occurring later on the same night.

Of the 618 incidents as possible precursors to 113 crimes, 18 incidents were found to be in the target 3 hour block before the crime and within 500 metres of the particular crime – an intersection rate of slightly below 3% (2.9%). This is clearly a low rate and suggests that there is significant spatial variation in the location of disorder incidents and assault crimes during an evening. A much higher intersection rate was expected as it was believed that incidents could be precursors to crimes later in the evening, and that an examination of the spatial interaction of the event types would show a correlation. This result demonstrates a very different outcome to the one expected and offers the opportunity to develop different hypothesis to explain the test outcome.

It may be that the attendance of the police at incidents in the early evening acts to diffuse the volatile nature of incidents and actually prevents the possibility of more serious events occurring in the same location later in the evening. Officers may be able to take the heat out of a difficult situation and commend the parties to a more peaceful resolution, or may threaten more serious authoritarian action if they are called on to return, and therefore the threat of arrest prevents further disorder. Once the police have attended a location the public know they are aware of the local problems and the perpetrators will be wary of attracting further police attention and more serious consequences. They will probably understand that if the local police have been unable to resolve a situation peacefully then they may resort to firmer measures.

Another possibility is that the police made an arrest (or number of arrests) at the early evening incident which removed a troublemaker from the scene and prevented the occurrence of a crime later in the evening. Arrests for disorder and assault tend to have a processing time at police stations measured in hours not minutes. The lack of cross reference between the incident and crime databases prevents the further exploration of this hypothesis. Even when arrests are not made, the incident may be serious enough that the local police management may decide that the area needs to more heavily policed and an immediate increase in police patrols and surveillance deters criminals from committing crimes in that location. Officers who attend the initial incident may take on this responsibility and remain in the area during the evening on the look out for more trouble. Active police patrolling for a specific problem may reduce the likelihood of a recurrence.

A less positive possibility must also be considered. It is possible that the individuals responsible for the committing of the crimes may be mobile and committing the crimes at different locations from the incidents, moving on to different sites throughout remainder of the evening.

The reason for the distinct separation in the animation is most likely a combination of the hypotheses in the previous two paragraphs. When the police are called on to deal with incidents they have a wide variety of sanctions and courses of action at their disposal. A desire to resolve the event at the first time of being called will mean that whatever the course of action, one police aim will be to avoid being called to the location again. The actions of the officer will reflect this and following a course of behaviour which will require another attendance on later occasions is unlikely.

It should be borne in mind that by narrowing the search criteria down to a distinct small temporal band for both incidents and crimes and requiring a correlation for a particular day, a large number of both data types were excluded from the analysis. This analysis is hampered by the lack of data and is reduced to using two data series which only cover a 90 day period. It is also vulnerable to the vagaries and complexities of crime distribution. It is also possible that problems build up over a longer time frame than a few hours before, and incidents a number of days before could be precursors of criminal activity. There is also the possibility of weekend to weekend crime, where incidents that happen one weekend indicate that the same location (for example a night-club) could be a 'flashpoint' on the next weekend.

To extract further information about the nature of incidents and crime data, more data is required. It will be especially interesting if more data becomes available to learn more about the suspects in incidents and crimes, and to accurately link the incidents which are the source call to recorded crimes.

5.8. CONCLUSION

This chapter has attempted to explore the temporal dimensions of crime and incident data using standard time series analysis techniques. The study area of Mansfield division in North Nottinghamshire was used as a test location for the techniques due to its extreme disorder and rowdyism problems identified in section 5.3 on page 110. With the two series in place, a number of processes identified both a weekend peak in crimes and incidents, and a difference in the temporal distribution of incidents and crimes throughout the day.

When the ratio of crimes to incidents is plotted on a map it becomes clear that the geographical distribution adds an extra complication to the temporal pattern. There is wide variety in both the ratio, and the distribution. This is the main revelation within this chapter. Few of the one kilometre grid cells in Figure 5-31 on page 144, or Figure 5-33 on page 146 display close to the expected ratio of values.

The attempt at correlating precursor incidents with reactive crimes had an extremely low intersection rate, and this is very interesting in both a temporal and spatial context. A number of hypothesis are proposed for this result, including the possibility that police involvement in the early stages of a dispute or incident in the early evening acts to prevent more serious transgressions of the law later in the evening. This could be either by arresting and removing an individual from the scene, or by diffusing the tension in the area. It is felt that the technique is worth retaining until a larger data sample covering a longer time period becomes available.

The wide distribution in incidents and crimes both spatially and temporally is all the more interesting when it is reiterated that the two data sets are not wholly independent. Without an appropriate cross-referencing procedure, identification of linked incidents was not possible.

The temporal and spatial analyses carried out during the chapter have revealed some of the complexity and difficulties in comparing recorded crime and incidents. More analysis with larger data sets may uncover additional aspects to this complicated relationship hitherto unrecorded and enable the predictive methodologies to be advanced.

6. Identifying repeat victimisation

The importance of repeat victimisation for crime prevention was highlighted in Chapter 2 (Previous work). This chapter's introduction describes the key literature that formed the background to this study, and identifies the areas of attention which were identified. The inclusion of georeferenced data in police crime records has created the possibility for rapid repeat location identification using the analytical power of a GIS. A methodology for this process is presented and a program structure defined.

The result of a burglary repeat victimisation study on a test area is then analysed. This study of burglary suggests that a standard GIS package, searching georeferenced crime locations can dramatically improve the time and accuracy of identifying repeats. The study employed two year's worth of data and the analysis raises issues regarding definitions of repeat victimisation.

This chapter therefore aims to use georeferenced crime data to identify repeat victimisation, apply the methodology to a study area, compare the time-course findings with previous work, and to examine the data in more depth in an attempt to identify 'true' repeats.

6.1. INTRODUCTION

The crime prevention benefits of accurately identifying repeat victimisation are undeniable. Chapter Two (*Previous work*) outlined the general background to the subjects of crime prevention and repeat victimisation and the gains that could be accrued if a system of rapidly identifying repeats were developed. It also described the manifest problems of identifying repeat victimisation locations from text-based police crime records.

Until recently most attempts at identifying repeat victimisation locations have focused on searching address fields in police records. Problems with inaccurate data entry and variation in address format make this method fraught with difficulty and time consuming to correct. The use of British Crime Survey results has highlighted the deficiencies in police crime records, and while the under-reporting of crime to the police is well documented (Hough and Lewis, 1989; Mayhew *et al.*, 1993), it remains a reality that the police recorded crime data is still one of the best sources of information on local crime distribution in this country. Computerised systems for recording police crime data have been set up within forces but usually the extraction of data pertinent to the geographical crime distribution and the identification of repeat victims is not a priority. A number of examples of this problem exist, and are quoted here to emphasis the technical background to the problem. For example (Ellingworth *et al.*, 1995):

> *All official sources of crime information are misleading. They uniformly fail to highlight the extent to which crime victimization is concentrated on particular individuals and households. (p.360)*

> *Police recording systems, in the author's experience, have been more or less inadequate in identifying repeated victimization of the same dwelling or the same people. (p.360)*

Other authors echo this point. In Huddersfield, David Anderson and his colleagues in the Huddersfield Criminology Group struggled for over four months to convert their data into an SPSS readable format.

> *The lesson of our experience with CIS [Crime Information System] for the project, and particularly for its evaluation, is that there is no easy way of measuring change in the proportion of repeat victimisations with any precision. (p.9)*
>
> *The problem of identifying repeats in police records is immense. (p.10) (Anderson et al., 1995)*

The Anderson study was not a specific geographical one. The geographical nature of the work was limited to the confining of search parameters to a certain divisional area, and the identification of repeat victims from a database of home addresses. This caused more problems as changes in beat boundaries during the course of the study limited the period over which repeats could be realistically calculated, and brought the repeat victimisation identification and study time down to a maximum of eleven months. Finally at the basic data level difficulties occurred with the manipulation techniques applied to the ASCII text data:

> *For instance, month fields were converted from alphanumeric to numeric, with JAN becoming 01, and so on. It was only later that we realised how often Black and Decker tools had been stolen. The process by which DEC had been converted to 12 had also converted these tools to Black and 12ker tools. (p. 6)*

With problems such as this, David Anderson and his colleagues are to be credited with having the will power to continue at all! Like other studies, they were rewarded with time course results that correlated with other work, and are discussed shortly. In a more recent study (Johnson *et al.*, 1997) the researchers had to write their own software to search for repeat victimisations. A Fortran program was constructed which combined a grid reference and address text-based search. Manual checking of the data for ambiguous results was a necessary final step.

These examples show the problems of trying to identify a unique location from a number of text fields within a crime data system. If search criteria are also tied to beat boundaries or other forms of containment then the problems of identifying repeats increase. Many of these problems arise due to the variety of ways in which a single address can be entered onto a database. As an example, the following fictional address could be entered legitimately onto a database in a variety of ways:

- 56 Easthill Avenue
- 56, East Hill Avenue
- 56 Easthill Ave.

These three occurrences of a single address would be recognised as different addresses by many computer-sorting processes. The problem is exaggerated when crimes are recorded on a computer by a number of different operators. Each operator may have a different understanding of Easthill Avenue, and will enter the data in a different format. Less experienced operators may also find difficulty in interpreting the handwriting of the police officers and enter a different concoction of words altogether. These problems arise when crime recording systems do not have modern gazetteers and in-built error checking routines that can confirm an address does exist, and can suggest corrections where necessary.

The difficulties multiply as the number of possible fields for an address increases. For example, Nottinghamshire Constabulary's Crime Recording Interim System (CRIS) has separate fields for: building name, building number, floor number, sub unit number, sub unit name, sub street name, street name, and so on. With such a variety of options it is hardly surprising that different operators on different shifts are occasionally likely to record a complex address in different fields from that of their colleagues. Nottinghamshire Constabulary has attempted to get round this problem by improving the address field data entry dialog on CRIS. The system will only accept addresses that it recognises, and then places the correct house number, street and postcode in the appropriate fields. This system can still fail, as it does not yet automatically enter flat numbers in shared buildings, or business addresses in small industrial estates that change regularly. Nevertheless this is a considerable improvement over a free-text entry system. Nottinghamshire Constabulary has gone to considerable effort and expense to improve CRIS, an option not available to all forces.

For a number of reasons studies into repeat victimisation tend to focus on burglary. While the type of crime is one which the public and police are keen to prevent and is often a police divisional and force priority, the underlying reason is often a more pragmatic one. It is relatively easy to spot burglary repeat victimisation as the location is often easier to identify than the victim. For the database search, the property, i.e. the location, becomes the victim. Anderson and his colleagues encountered definition problems trying to detect repeat victimisations outside the field of burglary (Anderson *et al.*, 1995):

> *A central concern of the project is the prevention of thefts of and from motor vehicles. Quantifying repeats here is intrinsically more complex, and merits a slight digression before the problems of CIS identification of vehicle crime repeats are described.*
>
> *A vehicle crime repeat could be defined in at least three ways:*
>
> - *Any vehicle, from the same complainant;*
> - *Any vehicle from the same location irrespective of its owner;*
> - *The same vehicle from any location even across changes of owner.*

The problem therefore is to find a method of extracting repeat victimisation records accurately and quickly from a crime database which is known to have conflicts within the textual address fields. Initial accuracy is necessary to prevent a time consuming follow-up manual trawl through the data, and speed is important because the highest chance of a repeat is known to be in the first few weeks after an initial incident. On a practical level, all these factors must be addressed if police forces are to utilise the information and it must be seen to be worth the investment in time and effort.

6.1.1. The time course of repeat victimisation

In the results of the repeat victimisation analyses, most studies in the field drew similar conclusions. Anderson and his colleagues (1995) found that 40% of all repeats happen within a month of the preceding one, while Burquest (Burquest *et al*, 1992) found an even greater figure of 79% of revictimisations occurring within one month for school burglaries in Merseyside. The generally accepted pattern is of an initial high rate of repeat victimisation which decreases rapidly after the first six to eight weeks.

Many of the studies in the literature cover shorter periods of between six months and a year. Longer study periods raise questions regarding the interpretation of repeat victimisation. Repeats, by definition, have a relationship with the initial incident. This might be because the same burglar returns to the location, or the burglar informs associates that the address is particularly vulnerable. At what point therefore do repeat incidents become unconnected with the initial occurrence? Is the burglary just another one, which happens to be at an address where a previous incident once took place? Polvi (Polvi *et al*, 1991) conducted a four year study (one of the longest studies in the literature) but failed to address this point, finding 'repeats' which occurred over three years from the initial incident. The drop in risk decreased in the Polvi data from an initial high until six months after the initial incident at which point the level of risk returned to the same as for the rest of the study area, demonstrating the absence of any elevated risk after 6 to 7 months. This correlates with general findings from other studies (Anderson *et al*, 1995; Burquest *et al*, 1992; Farrell and Pease, 1993; Hope, 1995; Polvi *et al*, 1990; Spelman, 1995).

6.2. IDENTIFYING A REPEAT WITH GEOREFERENCED CO-ORDINATES

It was decided to use MapBasic, the programming language for MapInfo, to identify repeats. This was mainly because the georeferenced crime data was already in MapInfo and the program could be applied seamlessly to this data. It would be feasible to create the program in another application, such as Excel or Access, if MapInfo were not accessible. The program is called RLFinder (Repeat Location Finder) and has been written by the author. A flow diagram showing the program structure is given in Figure 6-1 on page 163. The program incorporates a considerable degree of error checking including:

- Checks to see if a table is open
- Confirmation that the table has at least two columns
- Ensures the x and y co-ordinates columns are entered, and that these columns have numerical values in either float or integer format
- Creates a copy of the original table and works on the copy. This preserves the original table.

The program sorts the table using the Easting and Northing columns as the sort criteria. If a unique row is discovered it is marked for deletion. In MapBasic the row is not deleted until the table is saved, though it is marked as deleted. This means that the next row can be back-checked with a deleted row to see if they are identical. This unusual feature of MapBasic and MapInfo is the key to the operation of the program.

The full code of the program is available on the accompanying CD-ROM. The uncompiled program text is called **RLFINDER.MB** and a compiled version of the routine can also be found on the CD, titled **RLFINDER.MBX**.

The following tables use small example data sets to show the input and output data to the process.

Table 6-1 Example burglary data set with Easting and Northing values shown.

CrimeNo	FORCEmajor	FROMdate	TOdate	FROMtime	TOtime	Postcode	Easting	Northing
464	9	19960720	19960721	1700	946	NG2 2NL	457164	338668
10701	28	19960705	19960706	1730	830	NG2 1JT	457297	338633
873	9	19960731	19960801	1900	810	NG2 2NL	457121	338649
1032	7	19960726	19960726	130	130	NG2 2LW	457089	338653
2078	9	19960629	19960701	1300	700	NG2 2HU	457266	338882
5659	28	19970115	19970116	1900	1000	NG2 2FW	457382	338245
5690	28	19970128	19970129	2030	730	NG2 2FT	457369	338548
6494	28	19970201	19970203	1500	1600	NG2 2FT	457364	338494
6549	28	19961107	19961107	1745	1950	NG2 2FT	457369	338548
2173	28	19960728	19960729	2030	600	NG2 2HU	457266	338882
6634	9	19960610	19960610	2300	2300	NG2 2FS	457195	338328
6935	28	19960716	19960716	1000	1600	NG2 2FB	457237	338307
5119	28	19970311	19970311	150	150	NG2 2FW	457267	338525
7392	9	19960526	19960526	348	348	NG2 1LG	457319	338149
12111	7	19960527	19960527	1145	1145	NG2 1JT	457297	338633

Table 6-1 shows 15 values from an example burglary data set, with values for the crime number, Force crime classification (FORCEMAJOR), date and time variables, the postcode of the burgled premises and the AddressPoint Easting and Northing. Once the program has been run, the repeated values can be identified (**red** in Table 6-2) and retained.

Table 6-2 Data set sorted by Easting and Northing with repeat locations identified in blue.

CrimeNo	FORCEmajor	FROMdate	TOdate	FROMtime	TOtime	Postcode	Easting	Northing
1032	7	19960726	19960726	130	130	NG2 2LW	457089	338653
873	9	19960731	19960801	1900	810	NG2 2NL	457121	338649
464	9	19960720	19960721	1700	946	NG2 2NL	457164	338668
6634	9	19960610	19960610	2300	2300	NG2 2FS	457195	338328
6935	28	19960716	19960716	1000	1600	NG2 2FB	457237	338307
2078	**9**	**19960629**	**19960701**	**1300**	**700**	**NG2 2HU**	**457266**	**338882**
2173	**28**	**19960728**	**19960729**	**2030**	**600**	**NG2 2HU**	**457266**	**338882**
5119	28	19970311	19970311	150	150	NG2 2FW	457267	338525
10701	**28**	**19960705**	**19960706**	**1730**	**830**	**NG2 1JT**	**457297**	**338633**
12111	**7**	**19960527**	**19960527**	**1145**	**1145**	**NG2 1JT**	**457297**	**338633**
7392	9	19960526	19960526	348	348	NG2 1LG	457319	338149
6494	28	19970201	19970203	1500	1600	NG2 2FT	457364	338494
5690	**28**	**19970128**	**19970129**	**2030**	**730**	**NG2 2FT**	**457369**	**338548**
6549	**28**	**19961107**	**19961107**	**1745**	**1950**	**NG2 2FT**	**457369**	**338548**
5659	28	19970115	19970116	1900	1000	NG2 2FW	457382	338245

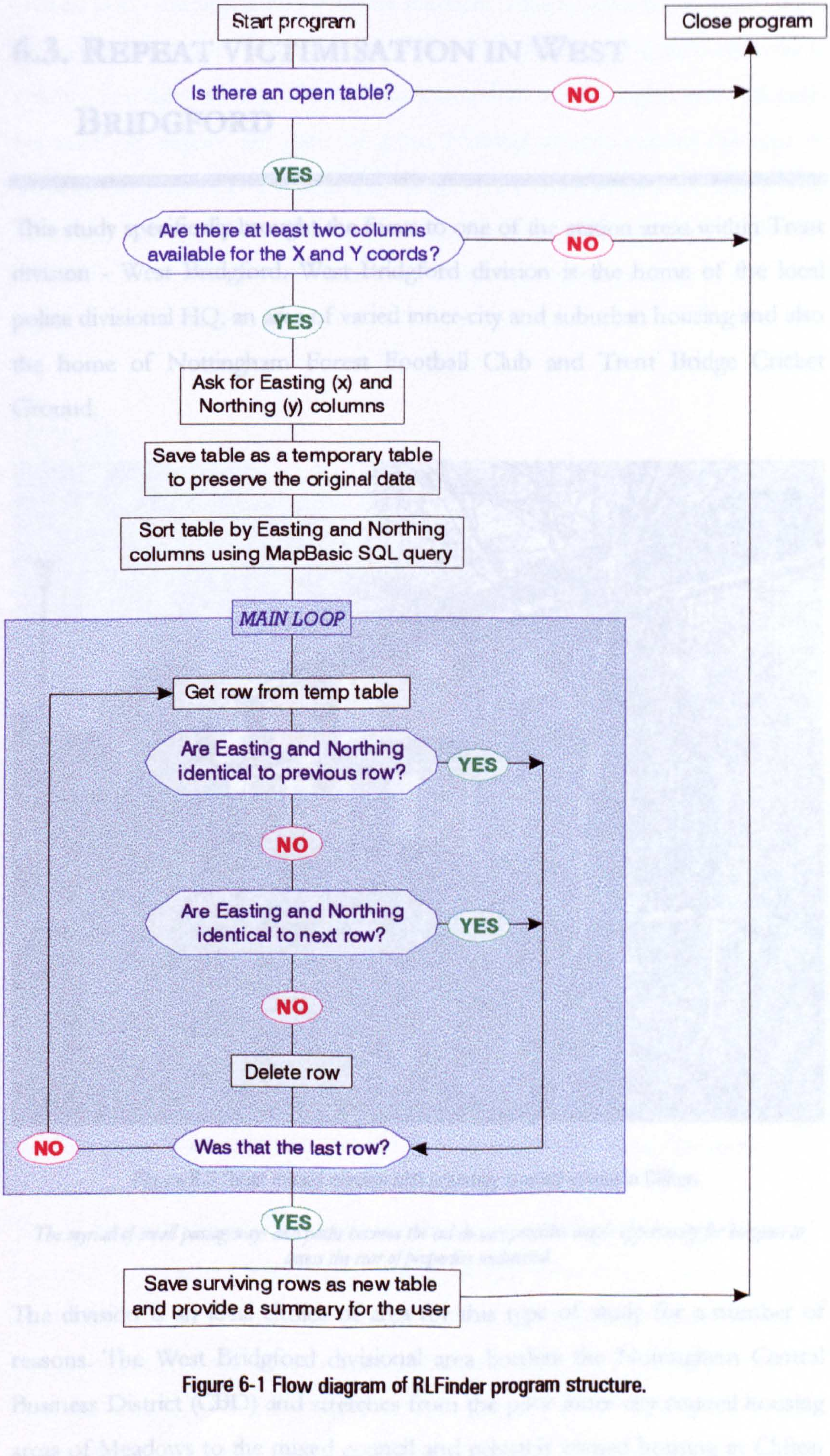


Figure 6-1 Flow diagram of RLFinder program structure.

6.3. Repeat victimisation in West Bridgford

This study specifically brought the focus to one of the station areas within Trent division - West Bridgford. West Bridgford division is the home of the local police divisional HQ, an area of varied inner-city and suburban housing and also the home of Nottingham Forest Football Club and Trent Bridge Cricket Ground.

Figure 6-2 Quiet mixed council and privately owned estate in Clifton.

The myriad of small passageways and paths between the cul-de-sacs provides ample opportunity for burglars to access the rear of properties undetected.

The division is an ideal choice of area for this type of study for a number of reasons. The West Bridgford divisional area borders the Nottingham Central Business District (CBD) and stretches from the poor inner-city council housing areas of Meadows to the mixed council and privately owned housing in Clifton

(Figure 6-2). The region also includes the leafy affluent suburbs of South West Bridgford (Figure 6-3). Each of these different types of housing and social mix is within a few miles of each other and they are often uneasy neighbours. Although the inner-city regions are short of space, Nottingham has resisted the urge to construct large numbers of tower blocks in West Bridgford division. This has an advantage for spatial studies as the identification of each property in two dimensions is made easier and individual burgled premises are easily identified as separate from their neighbours.

Figure 6-3 Leafy suburbs in the South of West Bridgford.

The South and Central West Bridgford area is considerably more affluent than the neighbouring regions of Meadows and Clifton..

The RLFinder program finds every location where more than one crime report has been recorded by selecting all records where there was at least one other record with an identical geographical location. For this particular example, repeat burglaries were chosen, but any other crime type could have been. The nature of these locations tended to be domestic properties with a few small light industrial estates also in the study area. The Nottinghamshire CRIS aims to attach to the incident a grid reference as close as possible to the *crime location*, not the grid location of the victim's home address (if they are different). Crimes that often happen in the street, such as assaults and personal robberies, are allocated the grid reference (via Address-Point) of the nearest building known to the system. Address-Point is aimed primarily at identifying households. The use of this type of data means that the police have difficulty recording the location of any offences which occur outside these 'known' locations, such as in the street or on waste ground. The one metre accuracy tends to place offences which occur in the street, in the kitchen (or at least the centre) of the nearest house!

At the time this study was completed (1997), Nottinghamshire Constabulary were in the process of changing from the Postcode Address File (PAF) to Address-Point data and some addresses in the county were still referenced in the crime data with PAF 100 metre resolution georeferences. It was therefore unfortunately necessary to check the data found by the GIS to remove erroneous records where the PAF grid reference was the same, but the address referred to different properties. It was heartening to discover that out of the dozen errors detected, only two were related to errors in the Address-Point data, the remainder being due to the resolution of the more inaccurate PAF.

6.3.1. Correcting for edge effects in the study time period

The data examined covered the whole of the 2 year time period, and a minor correction for edge effects was therefore necessary. In some studies which have looked at burglary figures, the researchers have been able to examine only a limited number of months data (see for example Anderson and colleagues, 1995, who examined 11 months of data for their study). If, for example, burglaries repeat uniformly within one month, the events that happen at the very beginning and very end of the study time period are at a statistical disadvantage. Events at the beginning are denied the possibility of being repeats to crimes that happened just before data recording commenced. For example, if a study time of one calendar year is used, events in the first week of January are denied the possibility of being a repeat of an incident in December because the December incidents are not included. Similarly crime events towards the end of the time frame are denied possible repeats in the next period. The current study, looking at crime figures over a two year period, showed that for the majority of the data set the greatest repeat time between burglary events to be 26 weeks (Figure 6-4).

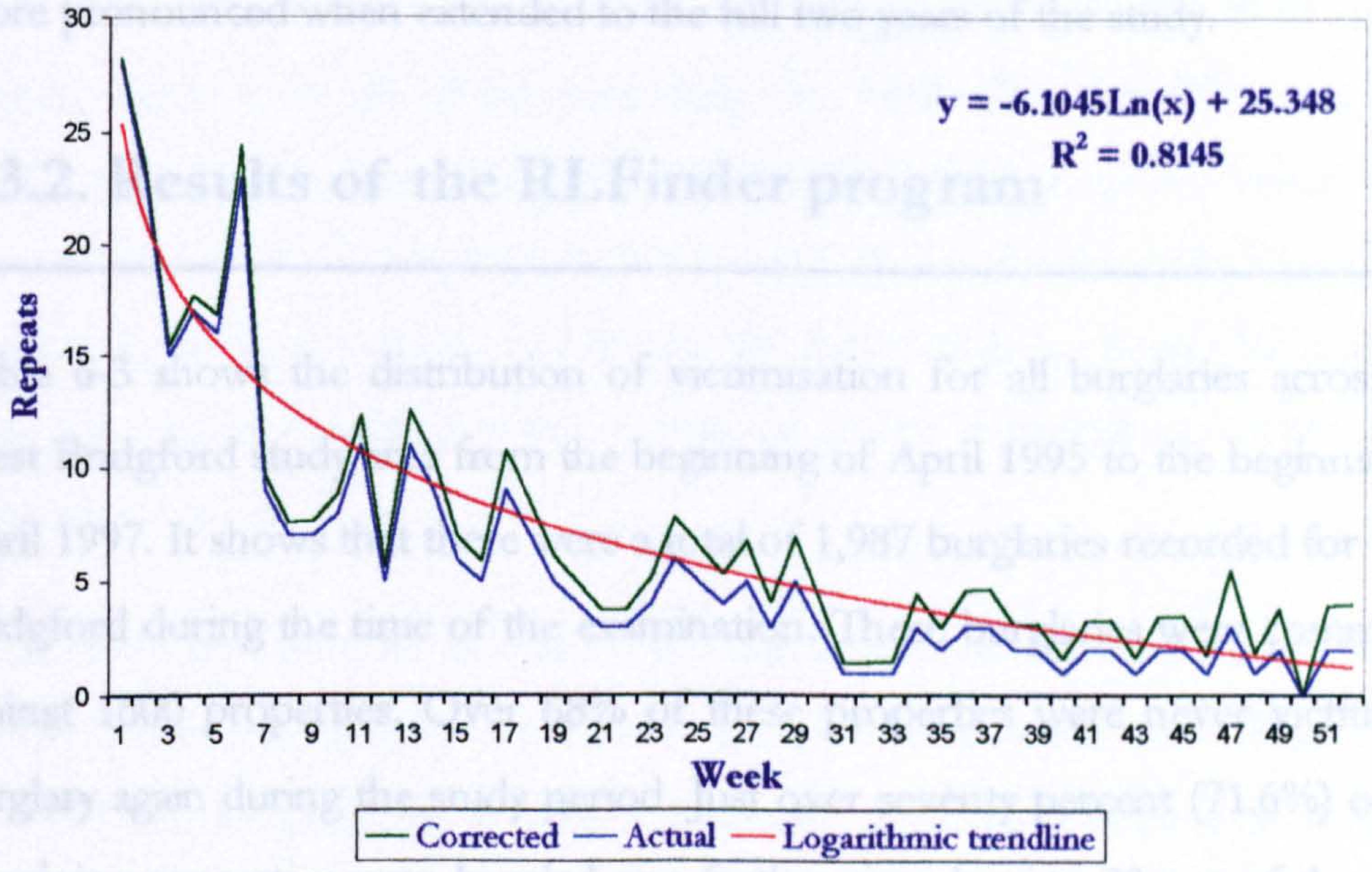


Figure 6-4 One year time-course of repeat burglary crime: April 1995 to April 1997.

Within the two year period (104 weeks), there were only 78 weeks where a repeat burglary half a year later would be included in the data available. If a repeat time of only two weeks is being recorded, there are clearly more weeks of data available which can contribute to the total. The correction used here is the same as employed elsewhere (Anderson *et al.*, 1995, p.47), and prevents the time-course graph declining artificially (Equation 6-1). The effect of this correction can be seen in Figure 6-4.

$$(\alpha\eta)/(\eta-\beta)$$ **Equation 6-1**

where: α is the number of events,

β is the number of weeks repeat time,

η is the total number of weeks in the survey (104)

There are problems with this formula related to the longer time frame used in this study which leads to over-compensation in the final months. The correction is more useful when restricted to shorter time scales. The effects of this over-compensation become evident as the x-axis values increase (Figure 6-4). Although the number of actual events drops below five, the number of incidents after correction with the formula is noticeably higher. This effect becomes even more pronounced when extended to the full two years of the study.

6.3.2. Results of the RLFinder program

Table 6-3 shows the distribution of victimisation for all burglaries across the West Bridgford study area from the beginning of April 1995 to the beginning of April 1997. It shows that there were a total of 1,987 burglaries recorded for West Bridgford during the time of the examination. These burglaries were committed against 1600 properties. Over 68% of these properties were never victims of burglary again during the study period. Just over seventy percent (71.6%) of the remaining properties were burgled one further time leaving 70 out of the 1600 properties victimised more than twice. Whilst these 70 properties account for only 3.5% of the number of victimised properties, they account for 14.3% of the number of burglary events. As in other studies (Johnson *et al*, 1997), burglaries

Table 6-3 Distribution of West Bridgford burglary victimisation: April 1995 to April 1997.

Times burgled	Properties affected	Incidents	Incidents (%)
1	1356	1356	68.2
2	174	348	17.5
3	39	117	5.9
4	11	44	2.2
5	8	40	2.0
6	7	42	2.1
7	2	14	0.7
8	2	16	0.8
9	0	0	0.0
10	1	10	0.5

which did not have a clear event date and could have occurred over a number of days, such as a weekend, were allocated the average of the starting and ending date. When the repeat time course graph is examined (Figure 6-4) it can be seen that the time between burglaries tends to agree with other published work in that the greatest risk of a repeat is the time immediately after a burglary. However this study found that the influence of the initial repeat period was lower than in other studies. Anderson and his colleagues (1995) found that 40% of all repeats happen within a month of the preceding one, while Burquest (Burquest *et al*, 1992) found an even greater figure of 79% of revictimisations occurring within one month for school burglaries in Merseyside. These compare with a much less dramatic figure in this study of 27.8% for the first month (up to and including 30 days) for repeats occurring within a year, or 23.3% if repeats up to two years apart are included.

There could be a number of reasons for this difference. Anderson (Anderson *et al*, 1995) examined domestic burglaries exclusively while Burquest (Burquest *et al*, 1992) was looking only at school burglaries. This study did not differentiate between types of burglary and examined all burglaries, domestic or otherwise. When the GIS was employed to select locations which had been revictimised twice or more frequently it was found that over 80% of the premises were non-residential. These included sports centres, schools and building sites. Police experience would suggest that thefts from building sites are inclined to have an extremely high reporting rate and that these crimes also tend to be reported as burglaries. Theft of tools from sites are reported by site foremen as the tools are often the financial liability of the contractors or have to be accounted for at a later stage. The loss of such tools is also reported preferentially as a burglary, as a theft might cast suspicion on work colleagues. Similarly schools and sports centres tend to have set guidelines for reporting burglaries and report a higher percentage of crimes that come to light.

The usefulness of georeferenced data to identify repeat victimisation is dependent on a number of factors. The problems already identified with the PAF would suggest that a database of crime data georeferenced with only PAF co-ordinates could be used only as an initial search tool. A manual search of the address data would still be necessary to remove erroneously juxtaposed crime events. Address-Point data, or similar unique address references would appear to be a substantial improvement on the PAF and would suggest that the identification of repeat victimisations within police data can now be greatly enhanced.

The use of georeferenced data would come into its own if the definition of repeat victimisation was enlarged to include larger geographical areas. Vehicle crimes such as theft from motor vehicles and theft of vehicles tend to occur in the street. Identifying a particular point on a street where a vehicle crime occurred from police recorded crime data is nearly impossible. The few studies that have identified repeat vehicle crimes have had difficulty in describing a repeat vehicle crime.

Areas such as street corners where a number of roads meet and are poorly-lit make parked vehicles an easy target. The crime reports for these events could show a number of different streets and would not be detected as a repeat by text-based search engines. The plotting of the georeferenced crime data within a GIS would enable the operator to identify the street corner as a possible source of the trouble. Although this is a case study of burglaries, motor vehicle crimes occurring at a street junction would be recorded as occurring at a nearby property and hence the power of a GIS is required to extract from the database all those vehicle incidents within some distance of the junction. This is one future direction for this type of work and enters the territory of 'hotspot' analysis.

Standard GIS procedures have been used in this study to locate repeat victimisation as an alternative to a macro written in a standard database language to identify locations that share identical georeferenced co-ordinates. The advantages of using a GIS over a standard database package are the ability to plot graphically the locations selected on a map for visualisation of the problem, combined with the ability to select repeats around a particular geographical feature, such as a road junction or town centre.

If a local authority wished to invest crime prevention resources in a geographical area such as a run-down estate, a text-based system would have immense difficulty in identifying repeats in that area alone. Such areas tend not to fall into particular police boundaries and a non-georeferenced search would either produce locations outside the desired search area, or worse, miss out locations within the estate. This is not a problem when a GIS is used as the user can more precisely define the area of the estate. The GIS can then be tasked to detect repeats in the specific streets of the estate and only those streets. It can also select the correct streets without fear of selecting crimes from identically named streets in other areas of the city or town.

6.3.3. Visualisation of repeats

A number of texts identify problems of integrating GIS and other technology into police forces (Hirschfield *et al.*, 1995a; Maltz *et al.*, 1991; Openshaw *et al.*, 1990). The use of graphical displays for viewing crime data makes the information more understandable and therefore accessible to often less technologically experienced police officers. The graphics make it immediately obvious where the majority of incidents are taking place on the sub-division. The examples given here shows how such displays have been employed in a typical software application (Ratcliffe and McCullagh, 1998a). Figure 6-5 shows a detail of the study area, focussed on a busy road junction in the centre of the suburban area.

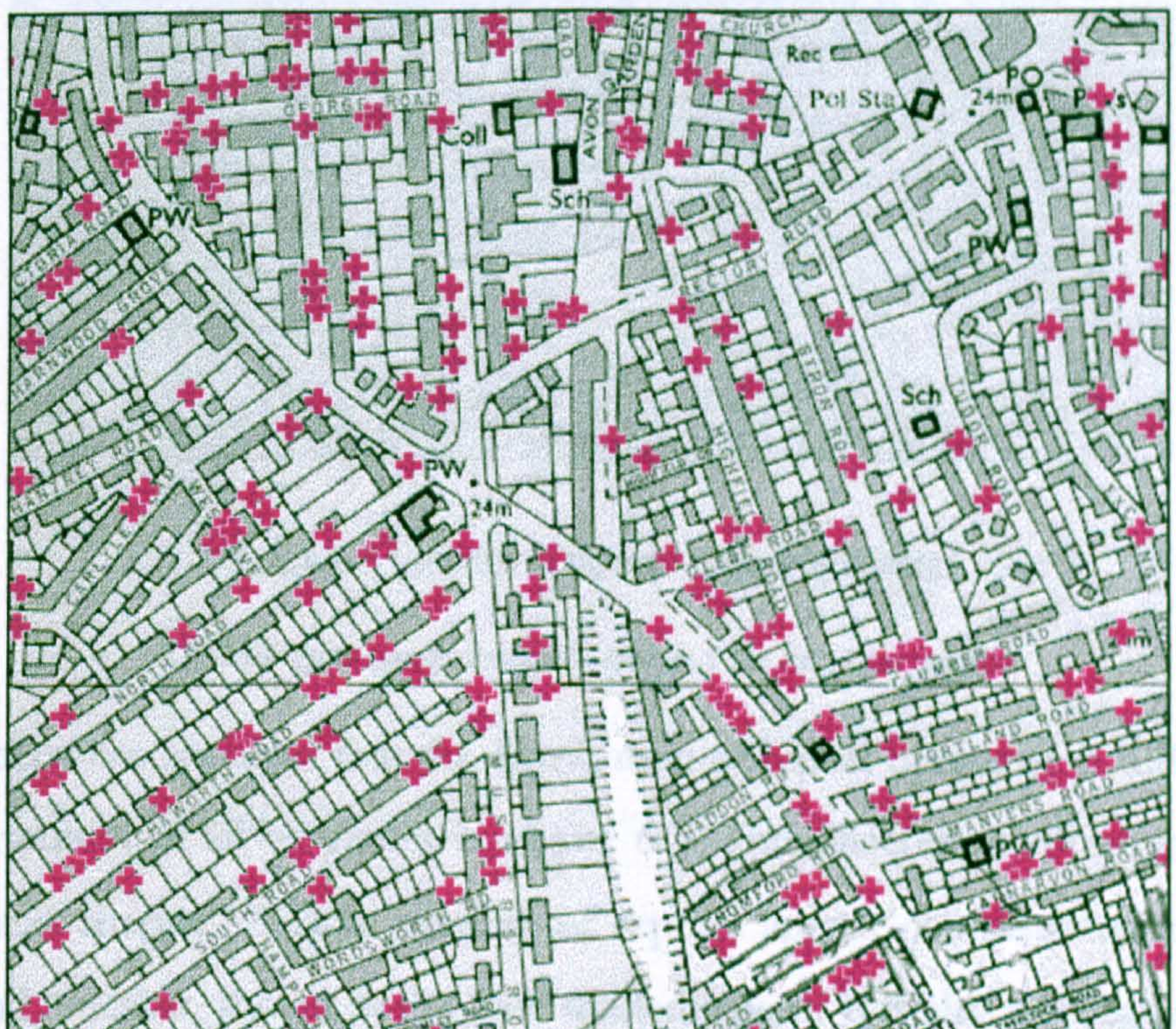


Figure 6-5 Location of all burglaries near a busy road junction (detail of the study area).

The purple crosses show the location of all burglary incidents over the two year period. The number of incidents at each location is not shown. If more than one incident happened at a location, only one cross is shown. The first step towards identifying repeat locations is to isolate the addresses that have been burgled more than once in the study period. Figure 6-6 shows the locations where a repeat victimisation occurred. The number of repeats is not shown, just the locations where the repeat events happened. Both images are portrayed over a raster 1:10,000 Ordnance Survey map. There is a high incidence of burglary across the study area, and the accurate targeting of crime prevention resources is certainly a local priority.

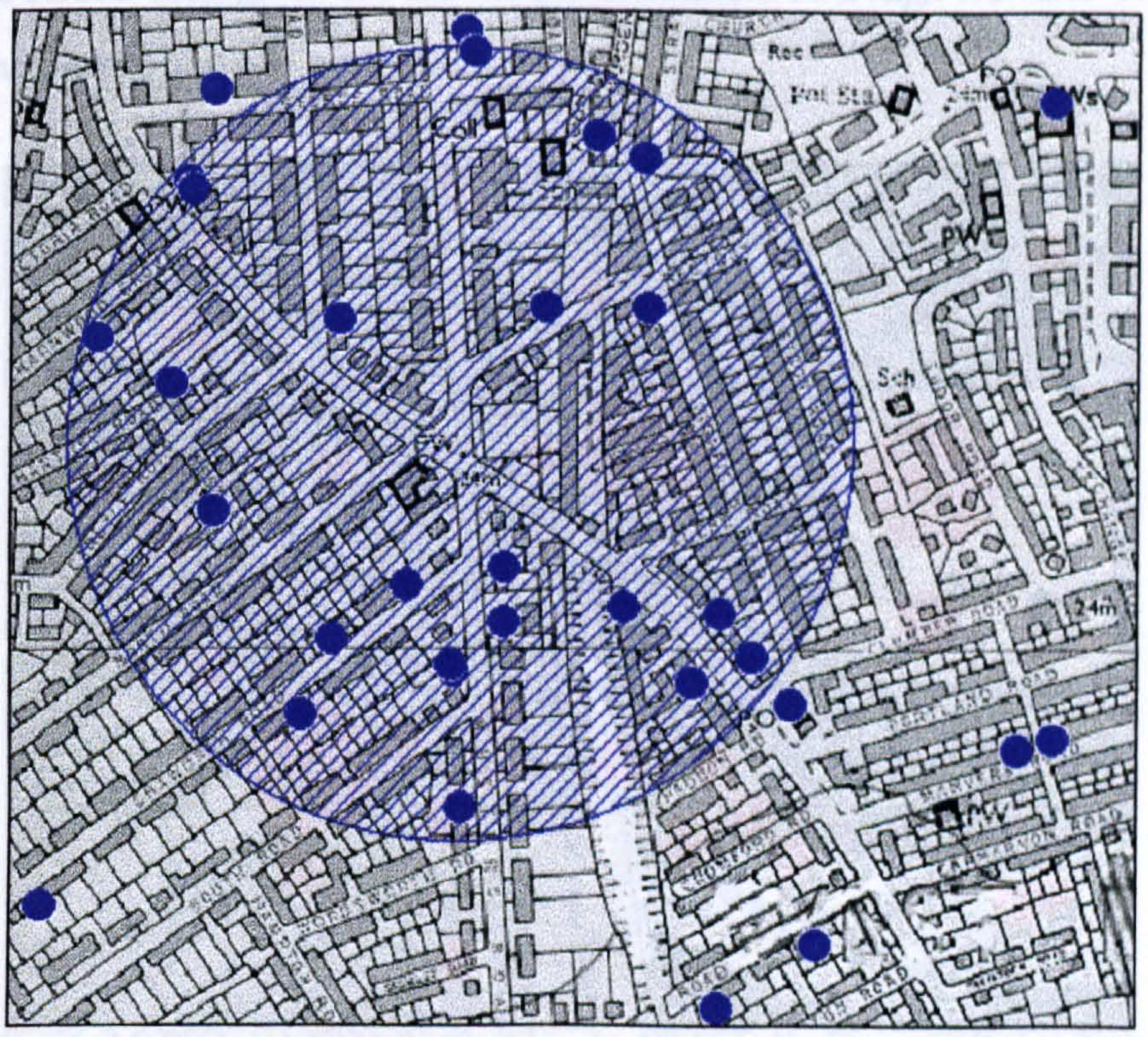


Figure 6-6 Location of repeat burglary incidents.

The blue region merely denotes the area near the busy road junction.

It then becomes apparent that the distribution of repeat burglaries is different to the distribution of all burglaries in Figure 6-5. This difference will be examined in the next chapter. A large number of repeats appear concentrated around a busy road junction, shown as the blue shaded region. It is also noticeable that in some areas of the map repeat burglaries are almost unknown. Possible reasons for this are that the residents took it upon themselves to improve their home security or other more complex social factors.

Figure 6-7 shows the same detail of the study area and the same locations where repeat burglaries have occurred. This map shows the magnitude of repeat burglaries at each address over the two year period. The larger the circle, the more crimes have been reported to the police as occurring at the location. One of the largest circles in the study (not shown in the figure) marks a council works depot that suffered 7 burglaries over the two years, and demonstrates the addressing problems discussed earlier. At the time of the study the depot was not included within the Nottinghamshire Constabulary address database and crimes at the location were assigned to the nearest property, a dwelling house in the same street. This gives the inaccurate indication of a high number of repeats at a domestic property.

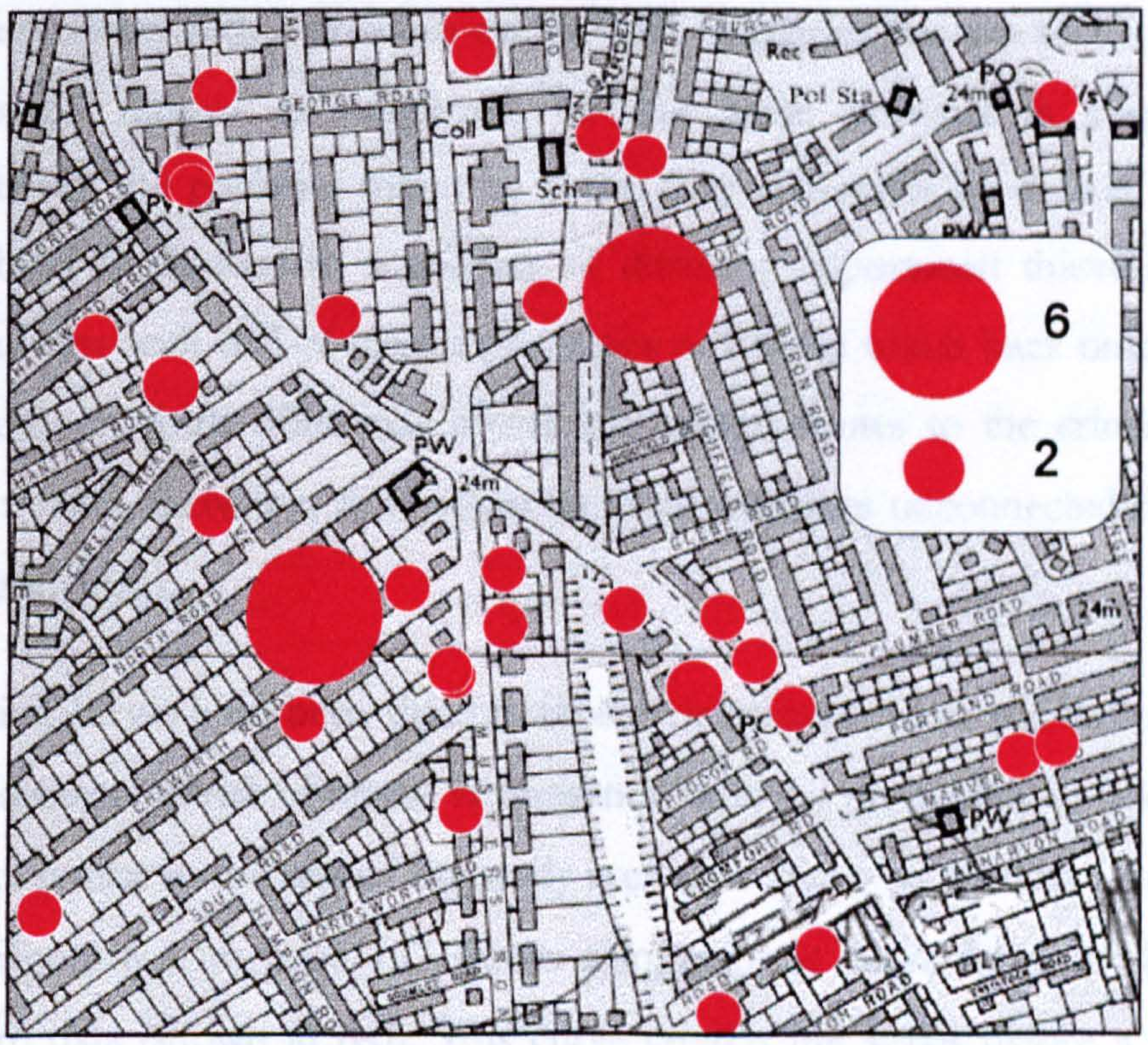


Figure 6-7 Magnitude of repeat burglary locations in the detail of the study region.

Larger circle size indicates a greater number of repeats at the location.

6.4. When is a repeat not a repeat?

Many of the other studies referred to here cover shorter examinations of repeat victimisation, between six months and a year, compared with this study. The longer two year study period here raises questions regarding the interpretation of repeat victimisation. The concept of a 'repeat' is that there is a connection of some description with a previous event. With burglaries this can be because the same *modus operandi* is employed, or the same criminal is responsible. Alternatively, the premises might be particularly vulnerable by default and this vulnerability is obvious to a number of different opportunist thieves. This is often the case with end-of-terrace properties or houses which back onto railway lines, both of which offer easy access and escape routes to the criminal. The question exists, therefore, as to when an event becomes unconnected from any previous incident.

Regression curves have been fitted to repeat victimisation data and these tend to show a decrease in risk of repeat victimisation until the probability of risk returns to a level similar to the rest of the study area after about six months (Polvi *et al*, 1991). Figure 6-4 on page 167 shows a logarithmic curve fitted to the West Bridgford data (shown in red). This curve crosses the x-axis before a year has passed, suggesting a similar decline in risk. A legitimate question is whether repeats that occur a year or more after the initial incident truly be considered *repeats*, or are they new *initial* events, unrelated to the original (or previous) event?

6.4.1. Similarities in burglary *Modus Operandi*

In an attempt to explore this question the original data was re-examined. Three variables were identified which might give some indication that a criminal or criminals committed the repeat incident with a similar *modus operandi* to the original event. Nottinghamshire Constabulary record the Point of Entry (POE) where the burglar gained entry to the location (for example; upper first floor sliding window, or roof skylight), the Method of Entry (MOE) employed (for example; using a glass cutter, or bodily pressure), and the times between which the offence was committed. This data is only recorded for burglary incidents, and is not fully recorded on many crime sheets. This means that this type of investigation is limited to burglary incidents and can not be applied to motor vehicle crimes unless the individual crime sheets are to be examined. The existence of this data for burglaries enabled a crude index of crime similarity to be constructed. Repeat incidents were considered matched if they had an identical POE, if they had an identical MOE, or could be identified as both being daytime (or night-time) burglaries. It soon became clear that errors or omissions in the data would seriously hamper the investigation. Table 6-4 shows the ratio of lost records. There were 631 incidents in the original repeat examination. Of these, 244 were 'originator' events, the first event at a property. This left 387 repeat events to compare with a previous incident.

Table 6-4 Quantity of records disregarded through incomplete data.

Reason for data removal	Number of records lost	Records remaining
STARTING POINT		387
No date record	28 (7.2%)	359
No accurate time record	5 (1.3%)	354
No POE/MOE record	94 (24.3%)	260
No time or POE/MOE record	24 (6.2%)	236

The reporting process is a definite source of error in this type of data. The officer at the scene reports the burglary and returns to the station where they fill in a crime sheet. The crime sheets are then entered onto the database by a civilian operator and not the individual who visited the scene of the crime. The system therefore relies on the reporting officer accurately describing the burglar's method and point of entry (if known). Absence of this information on the crime sheet usually results in an empty relevant field. These were detected and removed from the data, as shown in Table 6-4. Nottinghamshire Constabulary has tried to remove this source of error by redesigning the crime recording sheets. However there is also the problem of interpretation of the burglary crime scene by the officer. With 169 different POE options to choose from, it is likely that different officers will interpret and record the same POE differently from time to time. Other causes of the lost information come from communication problems within the force. If officers from another station record a crime, the details are telephoned or faxed through to the station responsible for the burgled location. If the officers are in a hurry, or the handwriting on the fax is not of a sufficient quality, some information can be lost.

COMPARISON OF REPEAT BURGLARY INCIDENTS

Initial analysis showed that from the 236 complete records in the data, 113 exhibited a match between occurrence and repeat in terms of either time of day, MOE or POE. It is assumed in this section that where a match exists, there is a reasonable probability that the repeat was directly related to the previous incident. If no match was found, as in 123 of the cases, this could suggest that the repeat was independent of the occurrence, and the apparent relationship between incident and repeat may not exist. A possible hypothesis to explain the lack of synchronicity is that the longer the repeat interval the higher the probability that a repeat is not a genuine repeat but a fresh incident.

Table 6-5 Ratio of similarity between matched and total 'repeats' over time.

Months between repeats	Ratio of similarity
1	0.75
2	0.61
3	0.59
4	0.60
5	0.56
6	0.45

There may be a number of reasons for a repeat not being a repeat beyond the criminal being a different individual. The environmental factors, which highlighted the property as vulnerable, may have changed through improved crime prevention measures or different policing patterns. This environmental change might be reflected in a different *modus operandi*. An attempt to investigate this possibility is shown in Table 6-5, where the ratio of matched repeats to total repeats for 30 day periods were calculated up to a six month limit. The similarity index ranges from 0 to 1.0, where a value of 1.0 indicates that all repeats have at least one aspect of commonality with their predecessors. In this data set the number of repetition burglaries becomes too small to be reliable beyond six months. However over the six month period there is a fairly steady decline in the proportion of matched repeats. This suggests that as lag-time increases the number of genuine repeats detected within the data declines. It must be borne in mind that this is an analysis based on a six month subset of the 236 'cleansed' repeat burglaries which occurred during the study time, and caution should be exercised when drawing conclusions based on the results of such a small study.

6.4.2. Testing the hypothesis further

A hypothesis from the previous section could be that the similarity between the repeat burglary and a previous incident decreases with time between the events. In support of this hypothesis, the analysis presented in the previous section shows a decline over the six month period.

The whole analysis was repeated in an attempt to see if the similarity decay phenomenon existed for larger data sets. Instead of using another sub-divisional area like West Bridgford, the exercise was repeated for two whole divisions of Nottinghamshire Constabulary. The process was applied to the burglary data from April 1995 to April 1997 for Beeston and Arnold divisions.

For Arnold division, 3741 burglaries were detected by the RLFinder program, of which 2557 actual repeats comprised the eventual data set of repeats (without their initial 'primer' events). Of these, 2338 burglary repeats took place within a year of the previous incident. At Beeston division the program found 1880 repeats of which there were 1254 actual repeats without their primer crime. From these, 1131 were found to have occurred within a year of the previous incident. As can be seen from these figures, a total of 3469 repeats for a year is a considerable improvement on less than 300 for a six month period.

The similarity index was calculated for Beeston and Arnold and is plotted alongside the West Bridgford data in Figure 6-8.

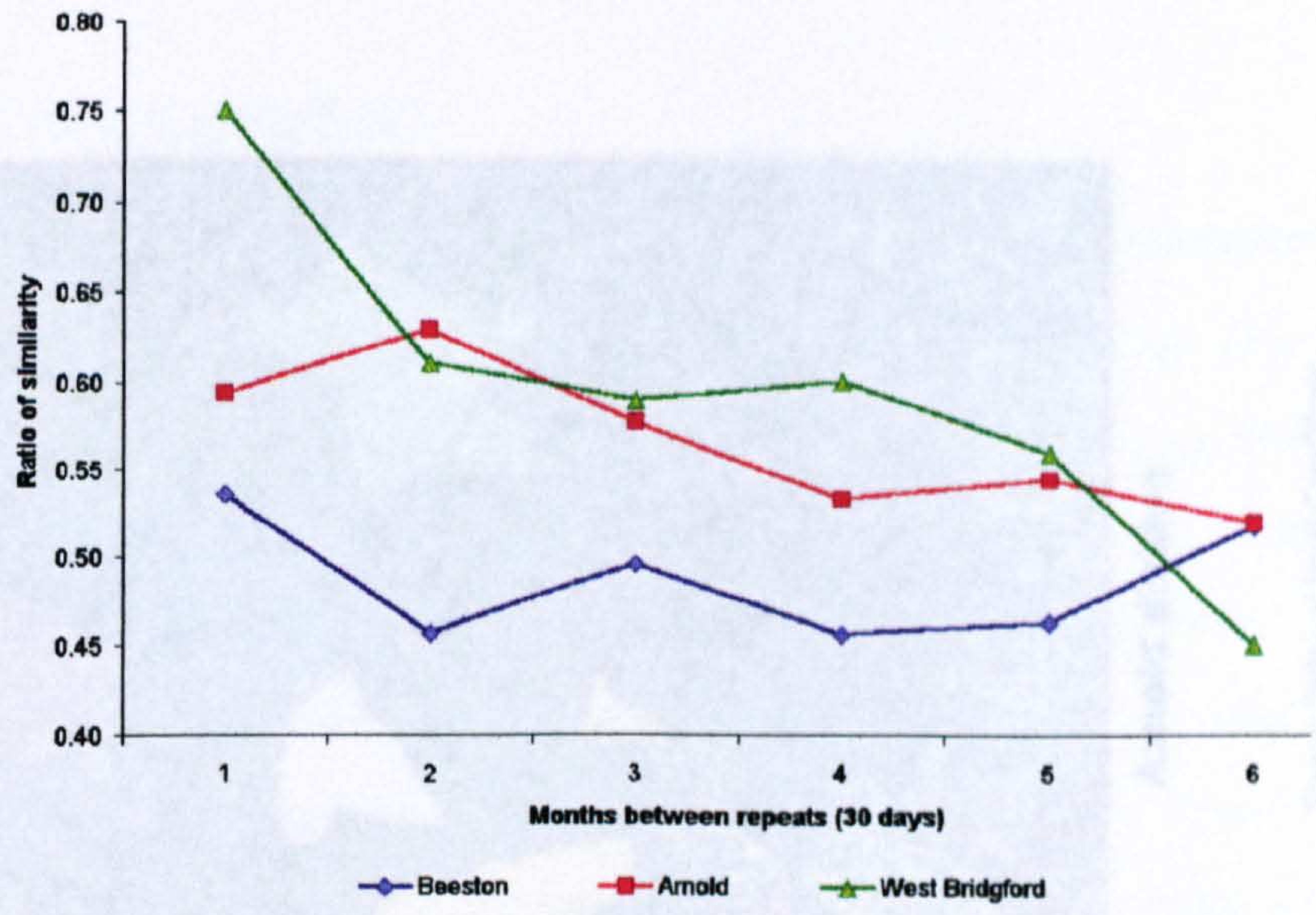


Figure 6-8 Similarity index for burglary repeats over 6 months.

The graph in Figure 6-8 shows that the West Bridgford data demonstrated the most noticeable decline in similarity as the time between burglary incidents increased. There is a similar but less dramatic decline in similarity for the Arnold divisional repeat burglaries and the Beeston data fluctuates around the 0.5 level.

There are a number of possible explanations for the differences between the data from the three test sites. The causes of crime are considerably complex and the search for a simple relationship between repeat time and crime similarity may be hampered by outside factors. The areas are relatively close geographically, but are quite different socio-economically. The architectural style, social fabric and economic prosperity of the regions are not only different between the police divisions but also show great diversity within their own boundaries. An example of the diversity between Arnold and Beeston division is shown in Figure 6-9. This shows the deprivation index values for enumeration districts in Beeston and Arnold police divisions. The more deprived areas of Arnold are concentrated in the South West of the division, while in Beeston the deprived districts are dispersed throughout the division. It may be possible that the socio-economic structure of an area has as great an influence on the propensity to repeat victimisation as temporal factors. This possibility is examined in the next chapter.

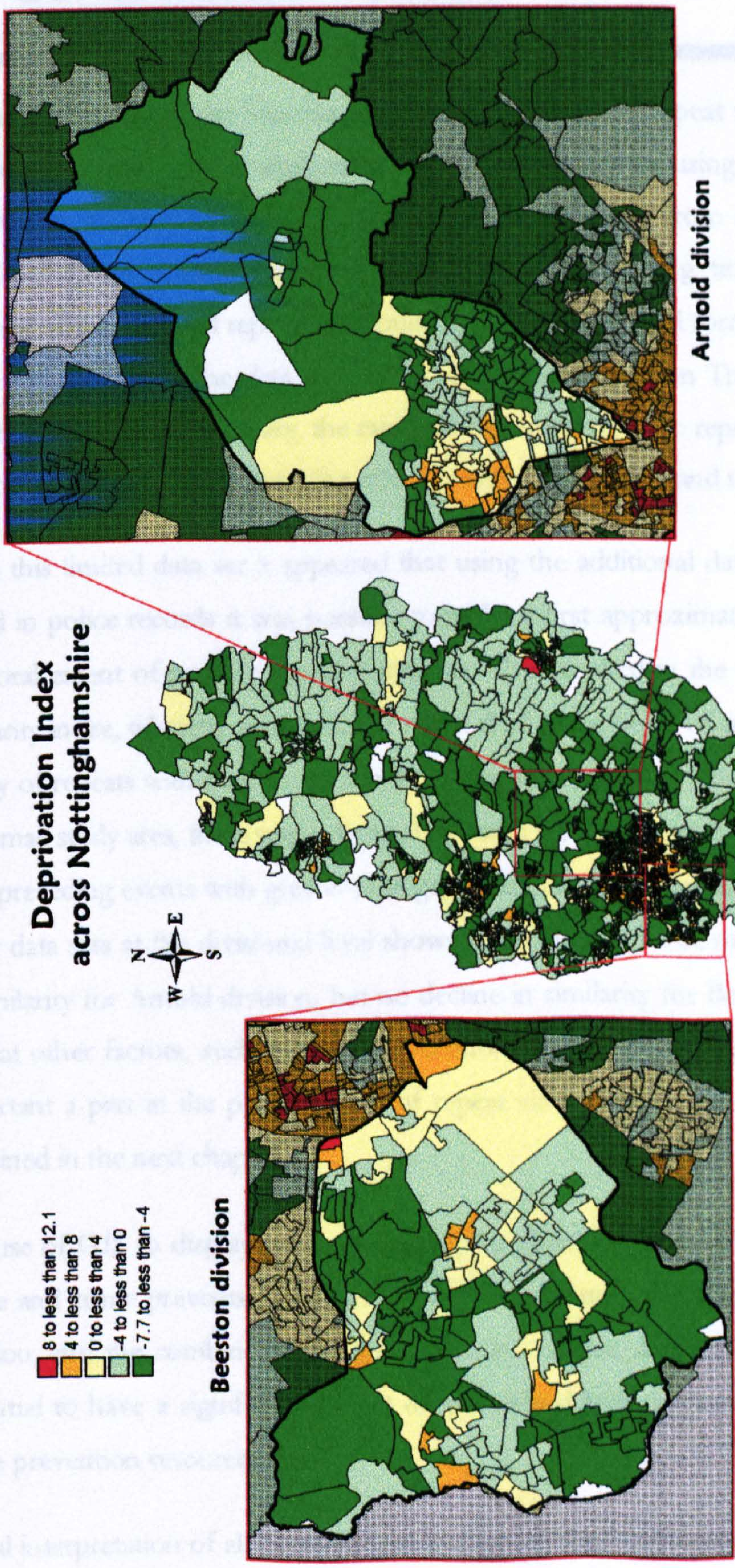


Figure 6-9 Deprivation index in Beeston and Arnold divisions.

Note that a positive value indicates greater relative deprivation

6.5. CONCLUSIONS

This study has shown the benefits of using a GIS to detect repeat victimisation. The accuracy and speed of analysis were greatly enhanced by using police crime data that included full Ordnance Survey geocoding directly from address data. This lead to great improvements in both overall processing time and final accuracy of the detected repeats. Although the repeats detected formed less than a third of the total crime data set for the study station area on Trent Division, Nottinghamshire Constabulary, the rapid identification of these repeats from the total crime data set is important for effective crime prevention and management.

From this limited data set it appeared that using the additional data commonly found in police records it was possible to make a first approximation as to the temporal extent of genuine repeat burglaries. This resulted in the creation of a similarity index, whereby similarities in method of entry, point of entry and time of day of repeats with their preceding incidents were recorded. On the evidence of a small study area, there was a decay effect with incidents losing similarity with their preceding events with greater time gaps. Further similar research applied to larger data sets at the divisional level showed a similar but more modest decline in similarity for Arnold division, but no decline in similarity for Beeston. It may be that other factors, such as the socio-economic fabric of the regions, play as important a part in the phenomenon of repeat victimisation. This possibility is examined in the next chapter.

The use of GIS to display the results of repeat victimisation searches can assist police and crime prevention authorities to assess the number of repeats at each location, and the combination of georeferenced address data and GIS has the potential to have a significant impact on police and local authority delivery of crime prevention resources.

Visual interpretation of all burglary sites and repeat burglary locations appears to suggest that repeat locations are differently distributed spatially from the general mass of burglaries, and this aspect of the data set is examined in the next chapter.

7. Burglary, victimisation and social deprivation

While several researchers (and the previous chapter) have noted the time course of repeat incidents, there has been little research that addresses the spatial variation in repeat incidents. In an attempt to explore and understand the differences between locations that are vulnerable to repeat attacks, and those sites that are victimised only once, this chapter uses an areally weighted approach to measure the level of social deprivation in the immediate vicinity of burgled locations. The chapter will show that this approach is important in avoiding the distributional problems which can occur if point data is aggregated to enumeration district level. A two year study shows that locations which are in deprived areas are significantly more likely to be the victims of repeat burglaries than affluent areas. A number of hypotheses for this phenomenon are discussed.

7.1. INTRODUCTION

The introduction to the previous chapter demonstrated how important an understanding of repeat victimisation is in the delivery of effective crime prevention (from such examples as Ellingworth *et al.*, 1995; Farrell and Pease, 1993; and Sampson and Phillips, 1995). Research has shown that targeting prevention at repeat victimisation locations can reduce crime, though rapid identification of the locations is important as victimisation tends to occur soon after previous events (Chenery *et al.*, 1997). Greater comprehension of the mechanisms of repeat victimisation is often hampered by difficulty in the extraction of repeats from the mass of crime data available to researchers and the police (Anderson *et al.*, 1995; Read and Oldfield, 1995; Sampson and Phillips, 1995), though as the previous chapter demonstrated, rapid retrieval of repeat locations is now possible with the aid of geographical information systems (GIS) (Ratcliffe and McCullagh, 1998a).

Most research has focussed on burglaries and these studies have identified a time-course in burglary repeat victimisation showing that the greatest risk of a repeat is in the time immediately following a burglary (Anderson *et al.*, 1995; Burquest *et al.*, 1992; Polvi *et al.*, 1991). This risk interval rapidly declines and after a few months returns to a hazard level similar to the general background rate. This repeat time course has received much of the research attention, but although the links between the general pattern of crime and social factors have been a popular area of investigation, the relationship between social patterns and repeat victimisation has been largely ignored.

Research linking general crime distribution and certain social conditions such as unemployment is common in the criminology literature (for a starting point see: Elliott and Ellingworth, 1996; Farrington, 1996; Hakim, 1982; Reilly and Witt, 1992). Relationships between deprivation and certain types of property and violent crime have also been identified (Herbert, 1976; Mayhew *et al.*, 1993). The link between crime and the social structure of an areal unit has received attention in Liverpool where both crime risk and proximity to underprivileged areas were examined (Hirschfield *et al.*, 1995), along with a principal component analysis of social deprivation and crime distribution (Hirschfield and Bowers, 1997). Like similar census-based studies this work aggregated crime locations by enumeration district and used the enumeration district boundary as the areal limit of examination. For a more detailed treatment of the available research on the link between crime and social factors, see chapter 2 (Previous work). Work completed in the previous chapter identified a possible spatial pattern after a visual interpretation of all burglary sites and repeat burglary locations. The apparent non-random pattern of repeats around a busy road junction appeared to suggest that repeat locations are differently distributed spatially from the general mass of burglaries. This aspect of the data set is examined in the current chapter, with the analysis extended to include the context of social deprivation.

This chapter also aims to improve the spatial interpretation of areal data around crime event locations and seeks a determination of the deprivation index value in the immediate vicinity of repeat victimisation sites. Locations linked by repeat domestic burglary victimisation are compared with points victimised by a lone burglary event against a background of social deprivation, to identify any differences in social fabric of the areas immediately surrounding the burgled premises. The emphasis in this study is on the individual location and not the number of incidents that have occurred at a particular site. In this way, it may be possible to elicit information about the vulnerability of buildings based upon the social deprivation context of the immediate area surrounding these insecure premises.

The existence of any relationship will be measured using an areally weighted spatial averaging method combining point data (crime locations) and the Department of the Environment (now Department of the Environment, Transport and Regions) Index of Local Conditions (ILC, ©Crown copyright) derived from census data associated with enumeration district boundaries. After demonstrating the technique on a small test area, using crime and ILC data from a suburban area south of Nottingham, the study is extended to a more heterogeneous area, enlarging the study region from the single suburban area to an entire police division with a broad rural/urban mix encompassing a wider variety of deprived and affluent areas.

7.1.1. The Vicinity program methodology

The purpose of the Vicinity program is to allow point data to be associated with an areal variable based on surrounding polygon data. Assigning a value to point data based on aggregation of census variables within polygons is a technique used when the count of points within the boundary is required for later comparison or mapping with associated census variables. The technique does not account for the proximity of a point to the edge of a boundary and the possibility that it could be extremely close to, and influenced by, another polygon within the boundary set. An extreme example of this problem would be a point internal to, but at the end of, a peninsula of one polygon, almost surrounded by a different polygon with significantly different census characteristics. The position of a point within a polygon becomes an important factor when calculating the data value to be assigned to the point on the basis of an aggregation of areal data surrounding the point. This is demonstrated in Figure 7-1a, where four fictitious enumeration districts contain a single crime point. The point (shown as the red dot) would be associated with the bottom right polygon, even though the left two polygons are close to the point. This is the peninsula effect.

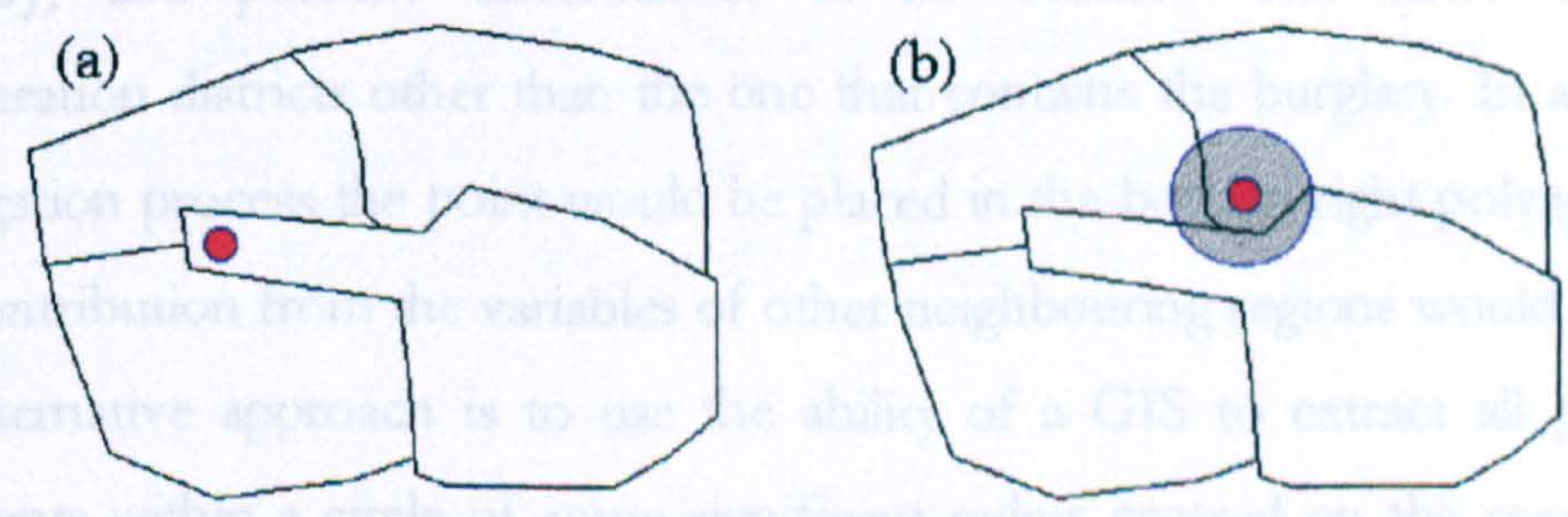


Figure 7-1 Peninsula effect and the possibility of error due to inaccurate georeferencing.

There also exists the possibility that a georeferenced point is not entirely accurate. Gatrell identified possible errors that can occur in the supposedly 100 metre accurate Postcode Address File (PAF). He also showed considerable errors can exist when the 100 metre resolution points are plotted in relation to their enumeration districts (Gatrell, 1989). This type of inaccuracy is demonstrated in Figure 7-1b where a different point (same enumeration district background) is shown within its area of error. The grey circle shows the region within which the actual location of the point might be. In the example this allows for the possibility of the true location existing in three different areas.

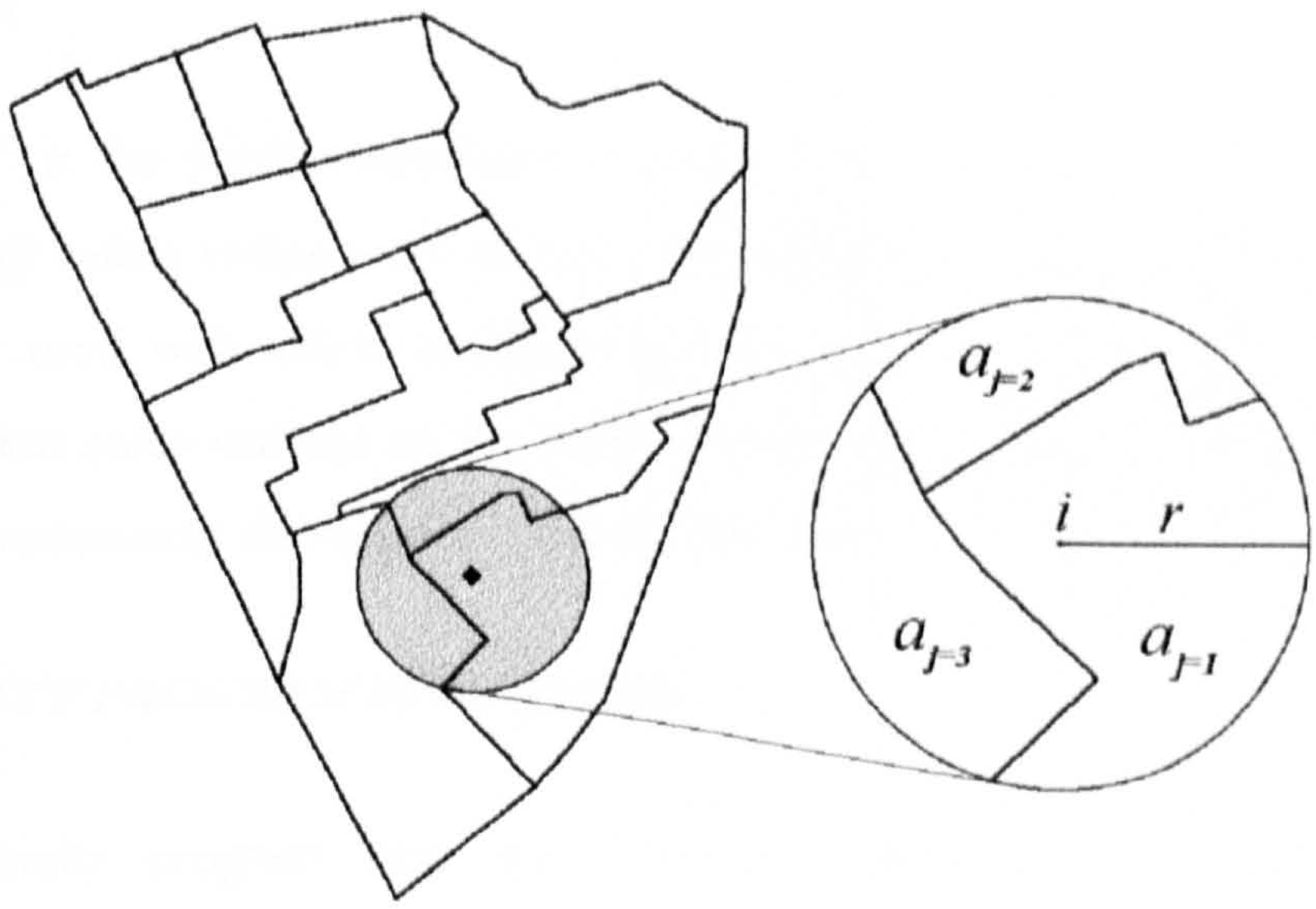


Figure 7-2 ***Vicinity*-type extraction of weighted variable from surrounding polygons.**

In this study the overlay capabilities of a GIS are used to isolate a region of areal data surrounding a point and extract the interpolated value of a variable based on its areally weighted contribution. Figure 7-2 shows a fictional ward showing 13 enumeration districts and a single burglary location, identified by the black diamond, together with a circle of size sufficient to reflect local influences on the burglary, and possible uncertainties in its location. The circle includes enumeration districts other than the one that contains the burglary. In a simple aggregation process the point would be placed in the bottom right polygon, and any contribution from the variables of other neighbouring regions would be lost. An alternative approach is to use the ability of a GIS to extract all polygon fragments within a circle of some significant radius centred on the crime. The parts of enumeration districts lying within the circle are isolated using a polygon overlay operation and a weighted average social deprivation index calculated using the values for each enumeration district weighted proportionally to the area that each district occupies in the circle as a whole. The final value of the variable for this location is based on the areally weighted average of the separate polygons within the circle (Figure 7-2), here called the "*Vicinity*" value (V) and calculated as:

$$V_i = \frac{\sum_{j=1}^{j=n} x_j a_j(r)}{\sum_{j=1}^{j=n} a_j(r)} \qquad \text{Equation 7-1}$$

where V_i is the *Vicinity* calculation of a variable at a point i, with a number of regions (j) within radius r of i, and $a_j(r)$ represents the area of a region within r of i. When used with crime locations and the ILC, we are able to calculate a deprivation index centred on the burglary location instead of using the location's single enumeration district social deprivation value.

VICINITY PROGRAM STRUCTURE

The Vicinity program uses the ability of MapInfo to alter and extract geographical regions from within polygons. The Vicinity program was written in MapBasic and automates a task which could be completed manually within MapInfo. The automation procedure is necessary when more than a handful of points are considered. Figure 7-3 shows the flow diagram of the program structure. Much of the program is of a robust construction with error-checking to permit the program to be used beyond crime analysis, and by less experienced users.

The program creates a separate table of circles after asking the user what size radius circles they wish to create. The choice of circle radius is an interesting question and is examined later in this chapter. The command line for this operation is:

```
Update circleTable Set Obj=CreateCircle

(circleTable.Easting,circleTable.Northing,radius)
```

As can be seen from this code, MapInfo is able to create geographical objects such as circles (command *CreateCircle*) and it is this type of operation which makes MapInfo an ideal tool for this type of program. The program is then able to enter the main loop of the program, the code of which is presented here (Table 7-1). Within this code, there are a number of purely geographical operations taking place. Each circle, centred on a crime location and with a set radius is used as the template for a 'cookie-cutter' operation (as in Figure 7-2).

Table 7-1 Key lines of code from Vicinity program main loop.

Line	Code	Explanation
1	`for rowNum=1 to totalRows`	Loop through each point circle
2	`select * from areaData`	Select the whole polygon study area
3	`Set Target On`	Make the selection the target for the 'cookie-cutter' operation
4	`Select circleTable.obj From circleTable Where RowID=rowNum`	Pick the next circle...
5	`Objects Intersect Into Target Data Vicinity_target=Vicinity_target`	'Cookie-cut' into the study area
6	`Objects Combine Data Vicinity_target=wtavg(Vicinity_target,Area(obj,"sq mi"))`	Combine regions into one using weighted average method
7	`tempValue = Selection.COL(areaVarCol)`	Extract the new combined value
8	`Update circleTable Set NewValue = tempValue where RowID=rowNum`	Save the value and associate it with the relevant circle
9	`Rollback table areaData`	Restore the area data table ready for next operation
10	`Next rowNum`	Next point circle

The area table has to be set as the MapInfo target for this operation (line 3) and then extraction of the polygon sections is performed. Once each circle is selected (line 4) the 'cookie-cut' (line 5) is combined (line 6) and the new value saved (line 8). The whole program structure is summarised in Figure 7-3.

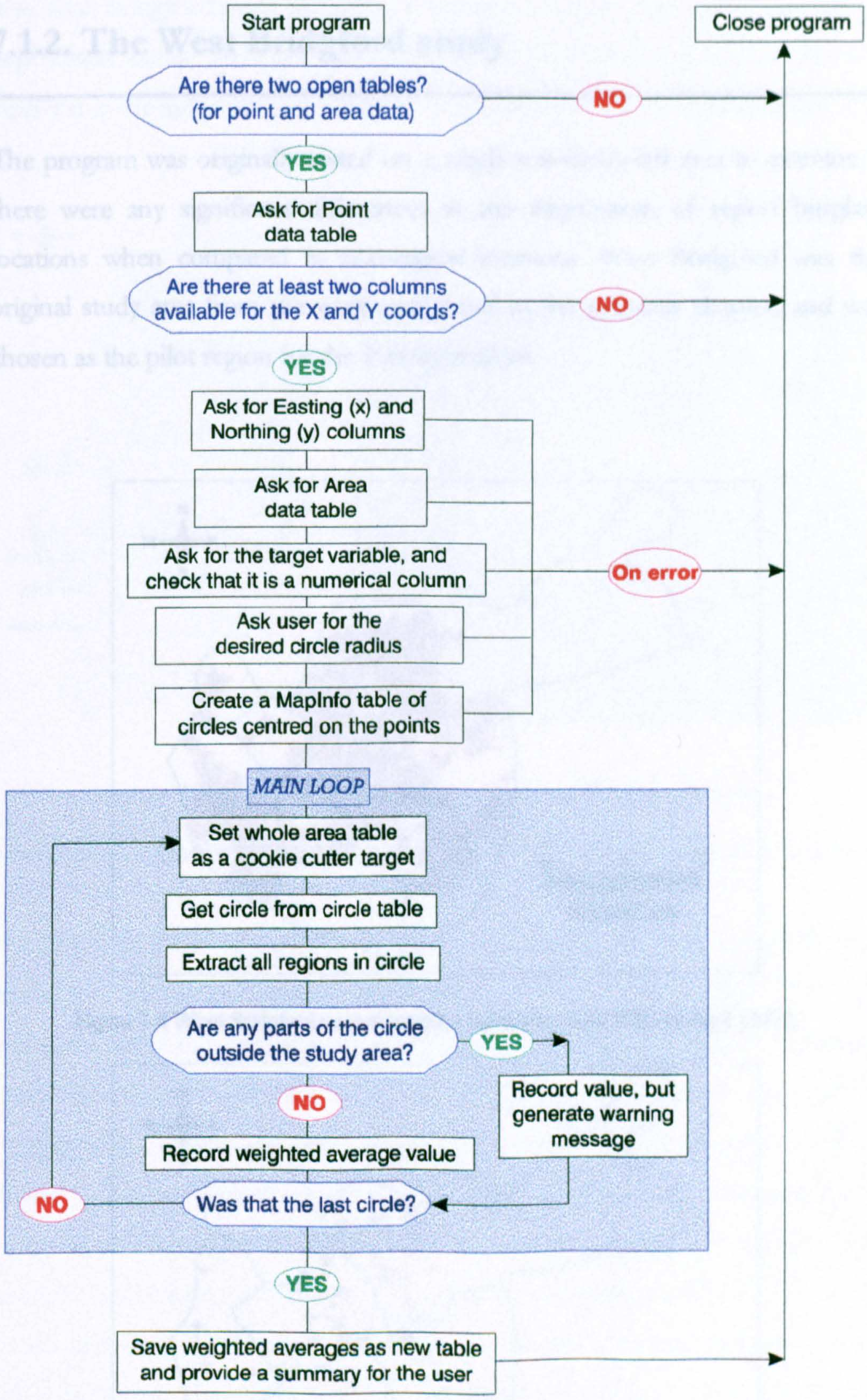


Figure 7-3 Flow diagram of Vicinity program structure.

7.1.2. The West Bridgford study

The program was originally tested on a single sub-divisional area to examine if there were any significant differences in the distribution of repeat burglary locations when compared to non-repeat locations. West Bridgford was the original study area from the work completed in the previous chapter, and was chosen as the pilot region for the *Vicinity* analysis.

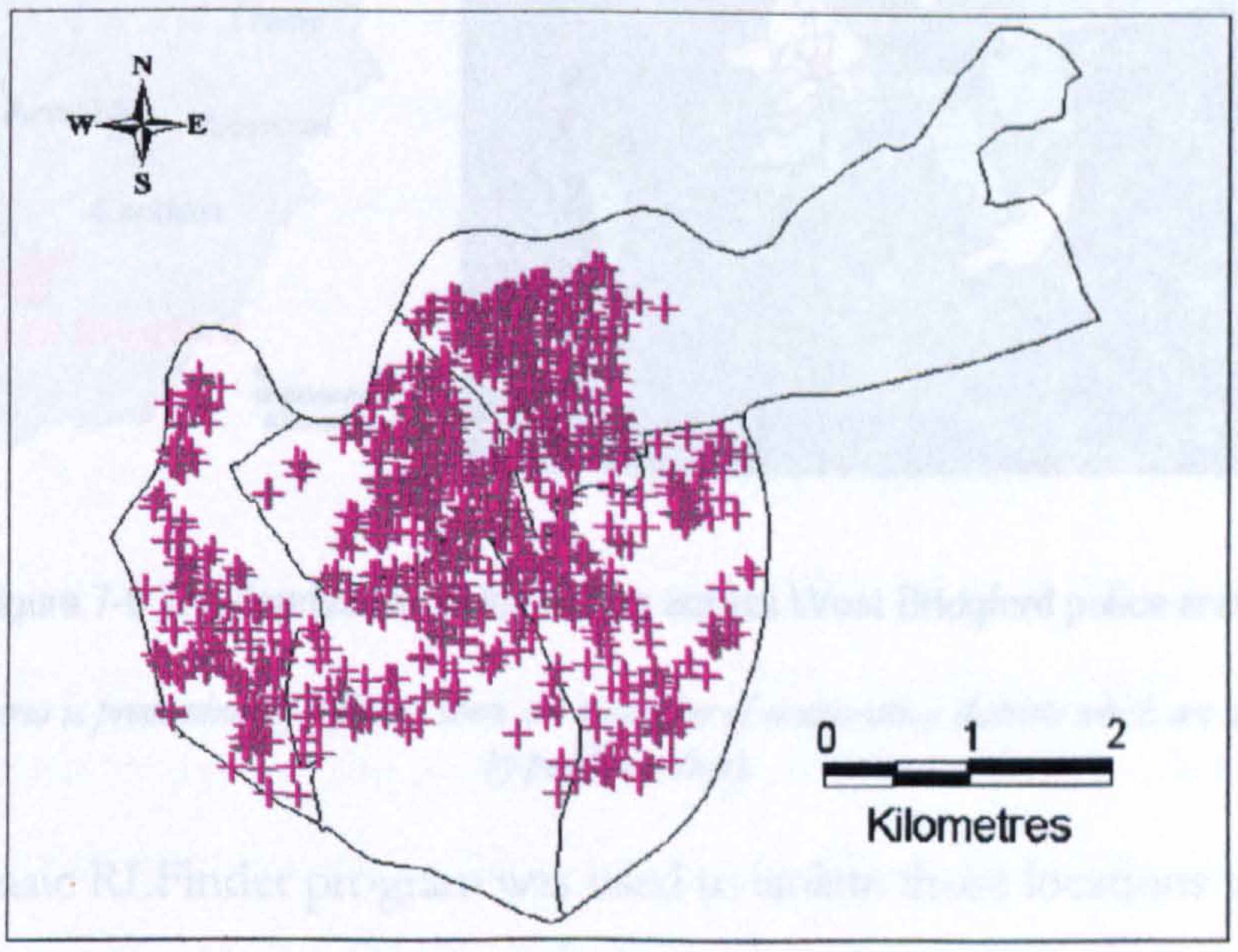


Figure 7-4 West Bridgford unique burglary locations (April 1995 to April 1997).

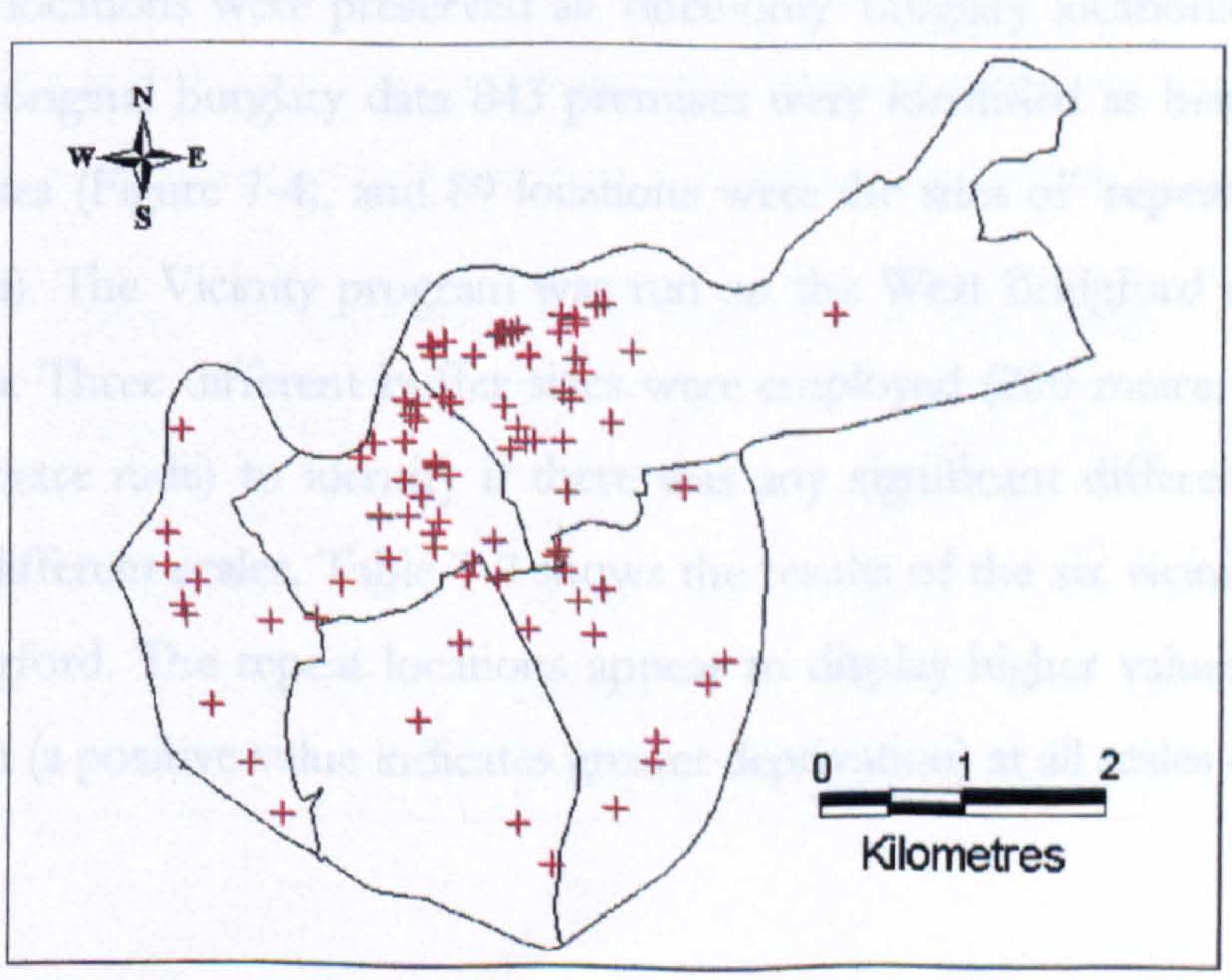


Figure 7-5 West Bridgford repeat burglary locations (April 1995 to April 1997).

The West Bridgford study area does not demonstrate as great a heterogeneity of deprivation index values as the whole of Trent division, but does show a variation in the index of local condition values in its mainly affluent suburbs.

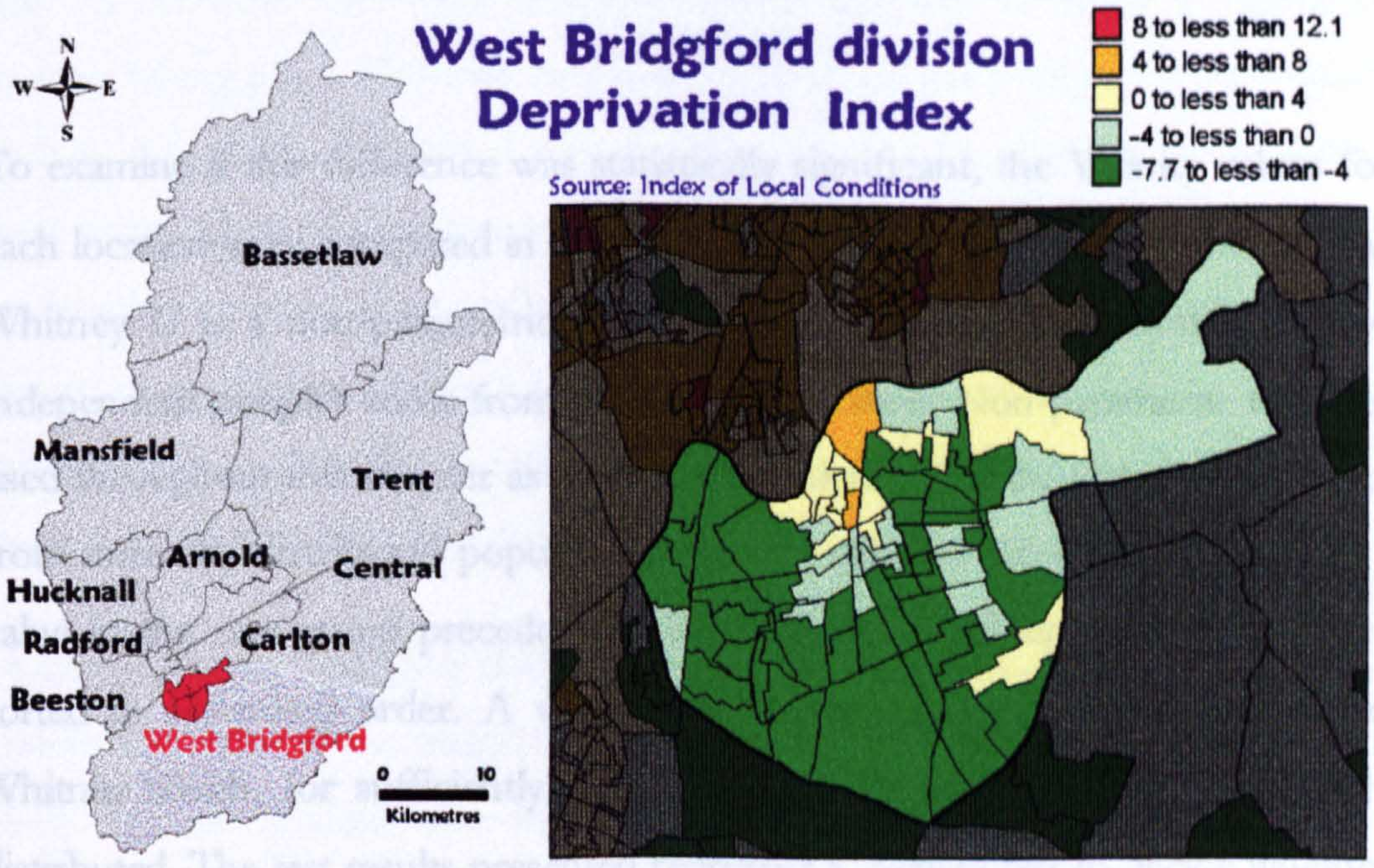


Figure 7-6 The deprivation index values across West Bridgford police area.

Although the area is predominately affluent, there are a number of enumeration districts which are deprived (shown by positive values).

The MapBasic RLFinder program was used to isolate those locations which were the target of more than one burglary in the period April 1995 to April 1997. The remaining locations were preserved as 'once-only' burglary locations (**unique**). From the original burglary data 843 premises were identified as being 'unique' burglary sites (Figure 7-4), and 89 locations were the sites of '**repeat**' incidents (Figure 7-5). The Vicinity program was run on the West Bridgford unique and repeat data. Three different buffer sizes were employed (200 metre, 350 metre and 500 metre radii) to identify if there was any significant differences in the results at different scales. Table 7-2 shows the results of the six vicinity runs for West Bridgford. The repeat locations appear to display higher values for social deprivation (a positive value indicates greater deprivation) at all scales examined.

Table 7-2 Mean values for six West Bridgford Vicinity calculations.

	200m buffer	350m buffer	500m buffer
Repeat locations	-1.869	-1.980	-2.051
Unique locations	-3.060	-3.080	-3.069

To examine if this difference was statistically significant, the Vicinity values for each location were compared in SPSS using a Mann-Whitney U test. The Mann-Whitney U is a non-parametric test to explore the null hypothesis that two independent samples come from the same population. Non-parametric tests are used throughout this chapter as there is no evidence that the sample data come from normally distributed populations. The U value is the number of times a value in the one group precedes a value in a second group, when values are sorted in ascending order. A value of z can be computed from the Mann-Whitney which, for sufficiently large sample sizes, is approximately normally distributed. The test results presented represent a 2-tailed test to decide whether the null hypothesis should be rejected. The value of p is the probability of obtaining results as extreme as the one observed, in either direction, when the null hypothesis is true. If the probability is small (.05 or less is often used), the null hypothesis is rejected. Table 7-3 shows that differences between the distribution of the vicinity values for the unique and repeat burglary locations are significant to at least the 0.005 level at all scales examined. The null hypothesis can therefore be rejected. Repeat burglary locations are in significantly more deprived areas than unique burglary sites across West Bridgford.

Table 7-3 West Bridgford Vicinity Mann-Whitney U test results.

	Unique (n=843)	Repeat (n=89)	Mann-Whitney U test		
Radius	Mean rank	Mean rank	U	z	p
200 metre	458.15	545.56	30477	-2.9134	0.0036
350 metre	457.20	554.61	29672	-3.247	0.0012
500 metre	456.07	565.26	28724	-3.639	0.0003

7.2. VICINITY STUDY OF TRENT DIVISION

The results from the West Bridgford study appeared to show a significant process taking place. However conclusions based on the results of a small study such as this should be treated with caution. The results, whilst very significant, were based on small data sets with 843 unique and 89 repeat cases. It would be preferable to extend the study both to increase the numbers in the data sets and to use a more heterogeneous area. This was done by extending the study to include repeat and unique burglary locations across the whole of Trent division to include affluent suburbs, council estates and rural villages.

Domestic burglary data was drawn from the force computerised crime recording system and covered the period April 1995 to April 1997. From this source data 3549 separate locations were identified as being the victim of a lone burglary during the study time, and 519 locations were identified as having at least one repeat incident. When mapped to the relevant Trent police division boundaries, 499 enumeration districts have boundaries within the study area. This data set includes the values from West Bridgford used in the previous study.

The question arose as to how big the integrating Vicinity circle should be for this type of analysis. The choice of radius for the analysis was considered relative to the average size of the enumeration districts in the study area. An upper limit could be a circle with a radius of about 750 metres as this would approximate the same area as the average enumeration district within Trent Division. Any larger and the process would begin to average too many areas and the individual characteristics of the enumeration districts would be lost.

There is a considerable variation in the size of enumeration districts in the Division owing to its mixed rural and urban nature. A smaller value would be essential in urban areas to reduce averaging of areally small enumeration district values, with possibly very different deprivation characteristics. The radius chosen must also be sufficiently large to ensure that any misplacement of the crime location due to standard recording difficulties is adequately allowed for; perhaps up to a distance of 100 metres. In addition in urban areas the radius needs to be large enough to allow for the peninsula location problem described previously. The process is designed to provide a reasonably continuous distribution of deprivation index values based on the values in the vicinity of the test location. Five different radii of 0, 100, 200, 350 and 500 metres (Table 7-4 shows summary statistics) were examined to see if distance affected the outcome of the analysis.

Table 7-4 Summary statistics of *Vicinity* analysis.

Buffer	Unique min.	Unique max.	Unique mean	Unique std. dev.	Repeat min.	Repeat max.	Repeat mean	Repeat std. dev.
0 metres	-7.51	12.09	-1.77	4.07	-7.51	12.09	0.24	4.64
100 metres	-7.51	12.09	-1.83	3.81	-7.51	12.08	0.17	4.39
200 metres	-7.51	12.09	-1.86	3.64	-7.51	11.62	0.03	4.12
350 metres	-7.21	11.93	-1.93	3.40	-7.24	11.39	-0.21	3.78
500 metres	-6.88	11.35	-1.99	3.19	-6.91	10.87	-0.42	3.50

Note that the analysis was conducted using all 3549 unique and 519 repeat burglary sites.

The 0 metres buffer returns the exact deprivation value for the polygon in which the crime lies, and was included to test whether the radii based *Vicinity* solution was generating vastly different and possibly unreasonable values. As expected the data extremes are reduced by the weighted averaging process compared with the exact values for a zero metre buffer, and are eroded steadily as radius increases (Table 7-4). At each of the chosen radius levels there is a noticeable difference in the mean calculations for the *Vicinity* results, with the mean levels for the repeat victimisation locations (repeats) appearing to be more positive (i.e. deprived) than for unique locations for all analysis scales.

7.2.1. Choice of Radius

The question remains as to which radius should be chosen to represent deprivation scores at a given crime location. A Kruskal-Wallis H test of all the unique burglary locations showed that there was no significant difference between radii choices in terms of calculated deprivation value. The chi-square value (χ^2) of 2.46 (approximation to the Kruskal-Wallis H) calculated from the five radius distributions based on the entire burglary data, with four degrees of freedom, and with p = 0.65, did not allow rejection of the null hypothesis of similarity between the deprivation statistics for different radii. Interestingly, when the test was repeated without the 0 metre radius data the calculated value of χ^2 dropped to 0.85 with p = 0.84. This indicated that the homogenising effect of the radius controlled areal weighted average calculation of deprivation generated very similar data sets to each other for different radii, but were marginally different from the 0 metre radius set. The reason for this difference lies in the spatial nature of the data which has at least as much importance as the statistical parameters of the data. Occam's Razor would suggest the 0 metre set should be preferred over the 100-500 metre circle sets, but the need to avoid the distributional problems of point locations mentioned earlier demands the acceptance of a spatially averaged smallest reasonable non zero radius tested – in this case at 100m.

7.2.2. Separation of Unique and Repeat Burglary Distributions

The spatial distribution of the unique burglary locations, and those at which two or more burglaries have been committed within the two year data set across the division is shown in Figure 7-7. Figure 7-7a shows the distribution of the 3549 unique burglary locations, and Figure 7-7b shows the distribution of the 519 sites which were burgled more than once between April 1995 and April 1997. The blue square delimits the area of magnification. As can be seen from these images, the majority of burglaries are concentrated in the areas of Trent division which border Nottingham city, and in the market town of Newark in the North of the division.

Table 7-5 Descriptive statistics of 5 different radii Mann-Whitney U tests.

	Group	n	Mean Rank	Sum of Ranks
0 metre	Unique	3549	1969.75	6990656.00
100 metre	Unique	3549	1966.04	6977462.50
200 metre	Unique	3549	1965.59	6975886.00
350 metre	Unique	3549	1965.30	6974840.00
500 metre	Unique	3549	1965.85	6976817.00
0 metre	Repeats	519	2477.24	1285690.00
100 metre	Repeats	519	2502.67	1298883.50
200 metre	Repeats	519	2505.70	1300460.00
350 metre	Repeats	519	2507.72	1301506.00
500 metre	Repeats	519	2503.91	1299529.00

A Mann Whitney U test was used to determine whether the population of the unique and repeat sets were significantly different (Figure 7-7). The descriptive statistics of the tests at all five radii can be found in Table 7-5. The analysis of the 100 metre data showed that the mean rank of the repeat victimisation locations of 2502.7 is consistently greater that the equivalent mean rank for the unique 'burgled once only' locations of 1966.0. The calculated Mann-Whitney U of 6777987 approximated to a z of -9.722 (Table 7-6 on page 199) which indicated that the null hypothesis of similarity could be rejected with considerable certainty ($p \approx 0.000$). These values are similar to the values of the 0, 200, 350 and 500 metre analysis (Table 7-6). It is concluded from this that the weighted deprivation index for the area in the immediate vicinity of Trent division repeat victimisation locations indicates significantly higher levels of deprivation than in the vicinity of unique burglary events.

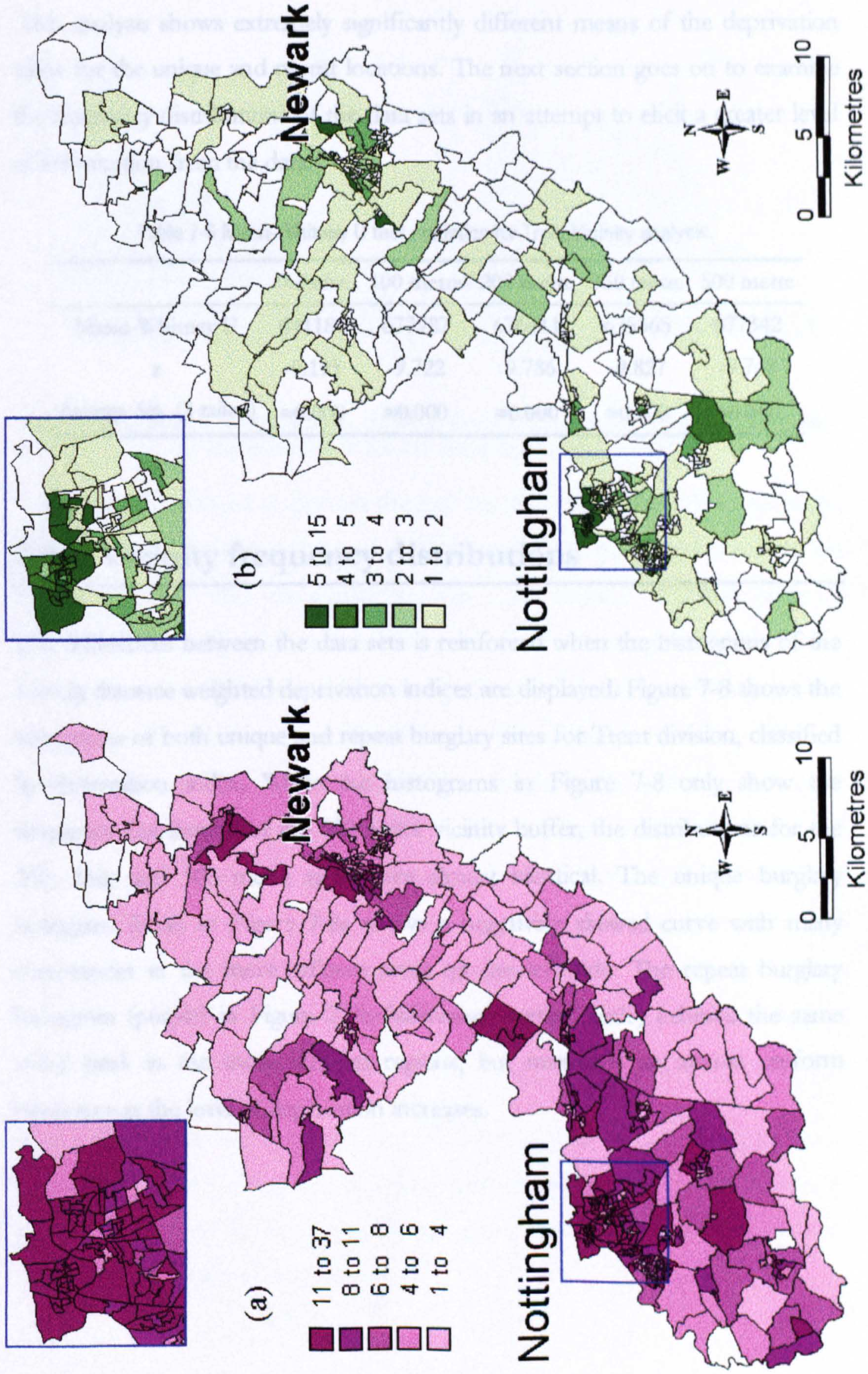


Figure 7-7 Trent division unique (a) and repeat (b) burglary locations.

This analysis shows extremely significantly different means of the deprivation value for the unique and repeat locations. The next section goes on to examine the frequency distributions of the data sets in an attempt to elicit a greater level of information from the data.

Table 7-6 Mann-Whitney U test statistics for Trent Vicinity analysis.

	0 metre	100 metre	200 metre	350 metre	500 metre
Mann-Whitney U	691181	677987	676411	675365	677342
z	-9.195	-9.722	-9.786	-9.827	-9.748
Asymp. Sig. (2-tailed)	≈0.000	≈0.000	≈0.000	≈0.000	≈0.000

7.2.3. Vicinity frequency distributions

The differences between the data sets is reinforced when the histograms of the *Vicinity* distance weighted deprivation indices are displayed. Figure 7-8 shows the histograms of both unique and repeat burglary sites for Trent division, classified by deprivation index. While the histograms in Figure 7-8 only show the frequency distribution of the 100 metre vicinity buffer, the distributions for the 200, 350, and 500 metre radii were almost identical. The unique burglary histogram (blue) in Figure 7-8a shows a negatively skewed curve with many occurrences in the more affluent areas on the left side. The repeat burglary histogram (purple) in Figure 7-8b (difference vertical scale) exhibits the same initial peak in the more affluent regions, but maintains an almost uniform frequency as the level of deprivation increases.

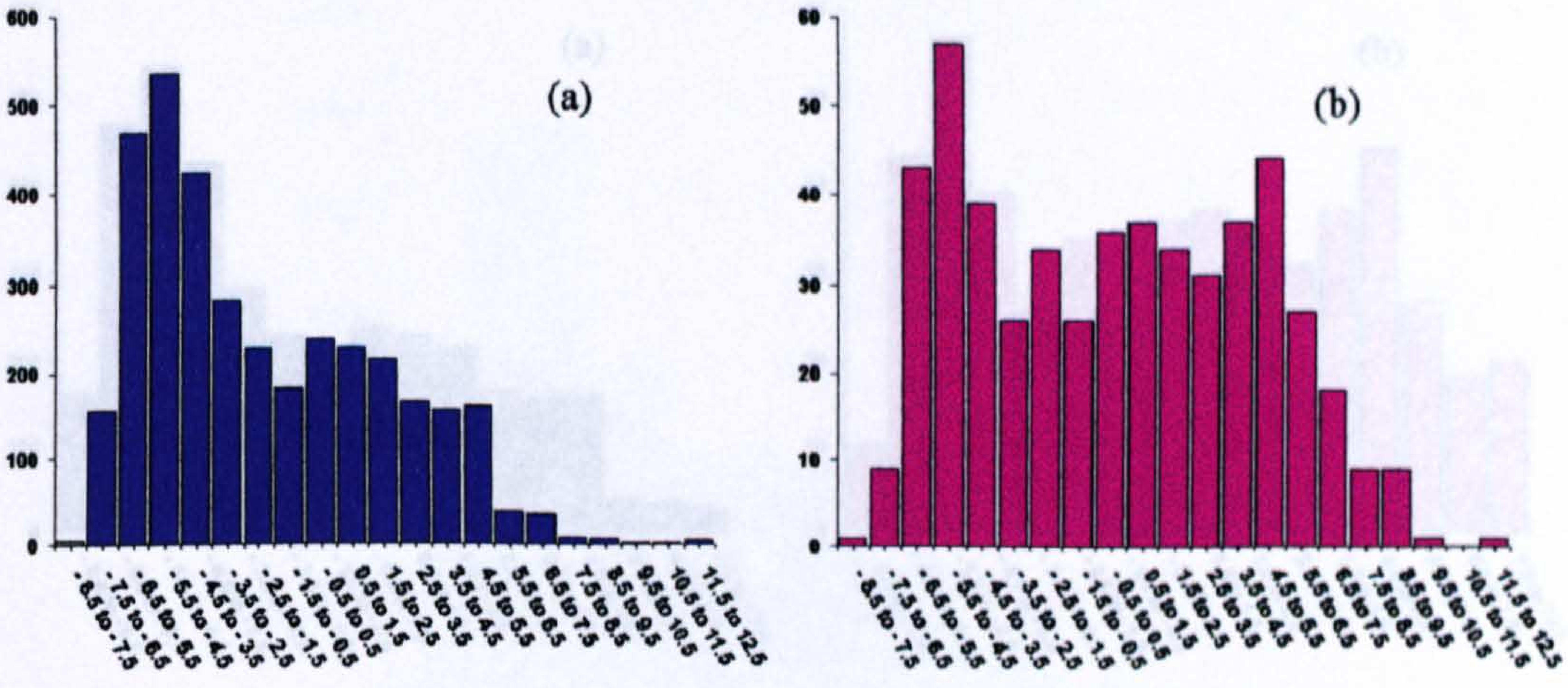


Figure 7-8 Histograms of Trent division unique (a) and repeat (b) burglary locations, categorised by deprivation index. The 100 metre Vicinity value is used.

A χ^2 test was employed to examine the null hypothesis that the differences in the frequency distributions in Figure 7-8 between the unique burglary and repeat burglary analyses were not significant, and a test statistic of 163.64 (with 15 degrees of freedom) was calculated. This involved collapsing two of the lower classes into one class, and also the top 5 classes into one, owing to low observed/expected numbers in those parts of the histograms (Ebdon, 1996). A restriction in this type of χ^2 test is that there should not be many categories for which the expected frequency is small. One rule that has been suggested is that if the number of categories is greater than 5 (as we have here), no more than 20% of the expected frequencies should be less than five, and none should be less than 1 (Ebdon, 1996). The effect of this transformation of the data can be seen in Figure 7-9 where the histograms show the actual frequency distributions which were used in the χ^2 test.

As the test statistic is greater than the tabled critical value of 37.7 at a significance level of 0.001, the null hypothesis is comprehensively rejected. The differences in the frequency distributions are significant and a measure of a real difference in the distributions of the two groups.

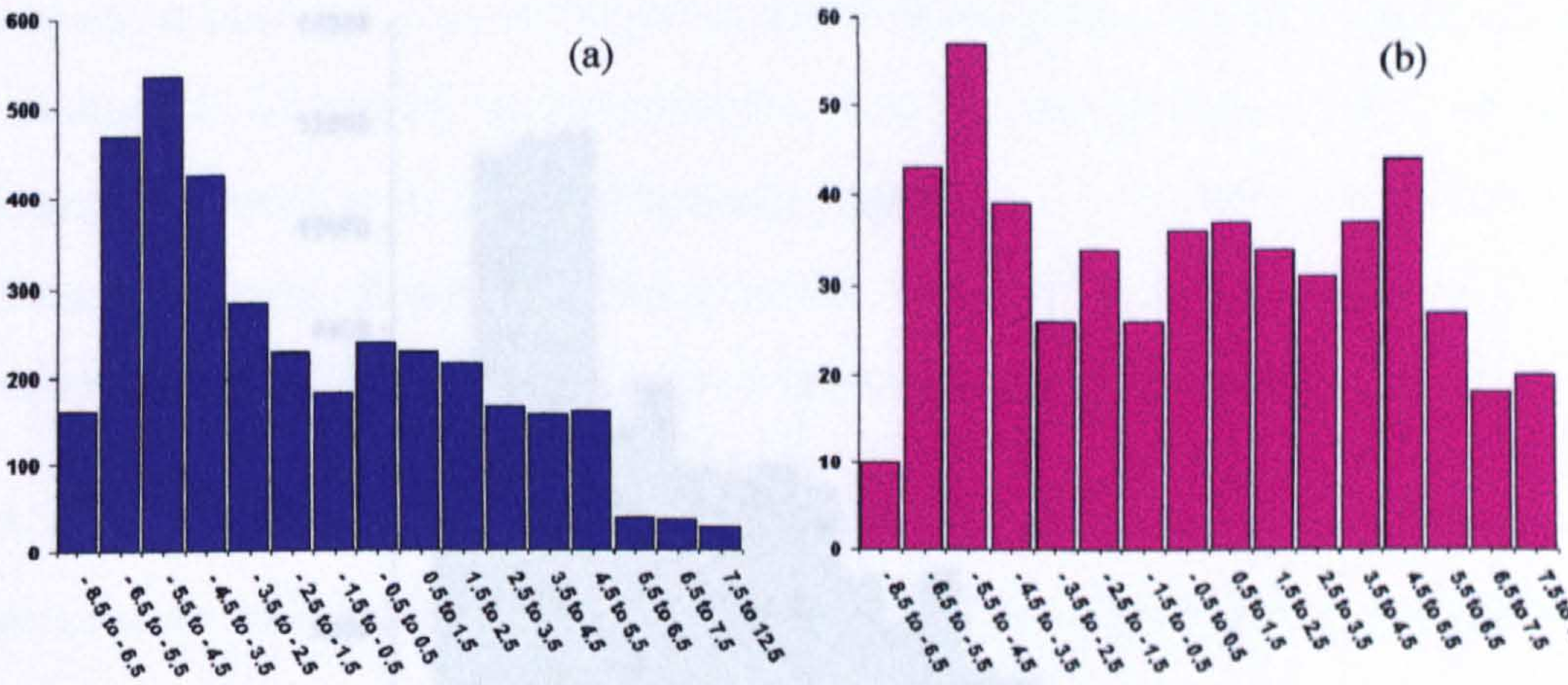


Figure 7-9 Compressed histograms of Trent division unique (a) and repeat (b) burglary locations, categorised by deprivation index.

There is therefore a significant difference in the distribution of repeat and unique burglaries in terms of enumeration district deprivation score within Trent Division. It would appear that the occurrence of repeat victimisation is found much more extensively in enumeration districts with a high deprivation index. The question remains as to which of the two distributions, unique or repeat burglary sites, more closely matches the distribution of the number of households by deprivation index. If the Vicinity values for unique burglary are more closely correlated to the household distribution, then this implies that the repeat locations are differently distributed and must be considered biased towards either affluent or deprived areas.

HOUSEHOLDS AND THE DEPRIVATION INDEX

The number of households in the Trent division area was calculated from the comma separated values files provided with the Map91 software package. This information was mapped to the enumeration district boundaries, and when compared to the deprivation index it was possible to construct a histogram showing the number of households in each category of deprivation index across Trent division. This is shown in Figure 7-10.

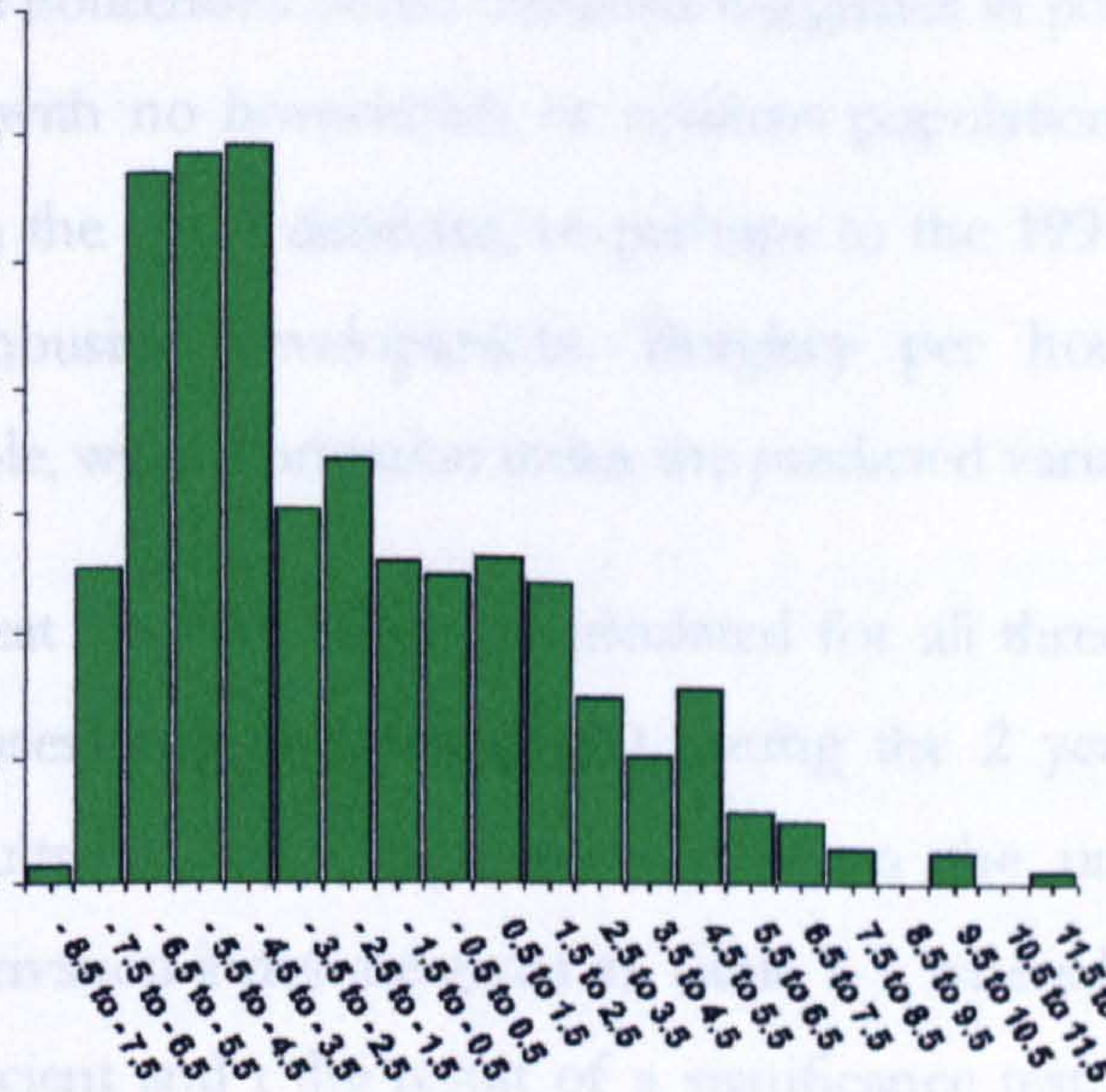


Figure 7-10 Trent division households and deprivation index histogram.

The graph shows the number of households across Trent division and their deprivation index category. A positive value indicates greater deprivation.

Figure 7-10 shows the distribution of households by deprivation index for the Trent Division. There is a notable visual similarity between this histogram and the distribution of unique burglary locations in Figure 7-8a on page 200, and Figure 7-9a on page 201. Whereas the fit between these two histograms is not exact it appears that the underlying model may well be similar, and quite different from that of repeat burglaries in Figure 7-8b.

7.2.4. Regression lines

A series of regression models were employed to test the significance of a linear regression fitted to the unique and repeat Vicinity models, and to look for significant differences between the slopes of the various regression lines. The data sets were derived from the enumeration districts (EDs) containing unique, repeat, and combined locations respectively. There was a considerable difference in the number of enumeration districts included in the three data sets from a low of 235 for the data set of EDs containing the location of at least one repeat burglary, through 439 for EDs where either at least one unique or one repeat (or both) took place, to 488 EDs that were found from the census information to include at least one household. Some domestic burglaries in police crime records occurred in EDs with no households or resident population – due either to locational errors in the crime database, or perhaps to the 1991 census data not containing new housing developments. Burglary per household was the independent variable, with deprivation index the predicted variable.

A unique and repeat burglary rate was calculated for all three data sets as the percentage of houses burgled in each ED during the 2 year crime data set window. The results of linear regressions between the unique and repeat burglaries and deprivation index are given in Table 7-7, where b_0 is the intercept, b_1 the slope coefficient and t the result of a significance test. For both unique and repeat burglary rates, the first null hypothesis tested is that the slope is not significantly different from zero (h_0: $*b_1 = 0$). The second null hypothesis is that the unique burglary rate (slope) coefficient is not significantly different from the repeat rate (h_0: $Ub_1 = Rb_1$).

Table 7-7 The relationship between the slope of a variety of linear regressions.[1]

Significance of slope, unique and repeat slope differences	Valid EDs [n = 488]			Unique EDs [n = 439]			Repeat EDs [n = 235]		
	b_0	b_1	t	b_0	b_1	t	b_0	b_1	t
Unique [h_0: $Ub_1=0$]	-2.71	0.15	2.70*	-3.87	0.34	5.82**	-3.61	0.38	4.94**
[h_0: $Ub_1=Rb_1$]			3.30**			7.26**			6.10**
Repeat [h_0: $Rb_1=0$]	-2.83	1.16	7.82**	-3.30	1.39	9.83**	-3.54	1.46	7.35**
[h_0: $Rb_1=Ub_1$]			27.23**			32.03**			24.77**

*Shown for different types of burglary, using burglary rates calculated per household to predict the Vicinity calculated deprivation index. Significance of t indicated by: * $p < 0.01$, ** $p < 0.001$.*

All the regressions were significant to at least the p = 0.01 level, and all the regression slope coefficients were significantly different from zero. All slopes showed a positive correlation between the number of crime events and an increase in deprivation index. In addition the slopes of the unique crime

[1] These results were generated using the regression curve estimation function of SPSS.

regression were always significantly different (the rejection of the null hypotheses of $Ub_1 = Rb_1$ and $Rb_1 = Ub_1$), and flatter than those for repeat burglary regressions for the same ED data set (shown by a lower value of b_1 for each of the three data sets). The first set, valid EDs in Table 7-7, showed the lowest slopes of the three sets because of the inclusion of many EDs where households existed, but showed either no burglary at all (47 cases), or alternatively no repeat burglary (253 cases). As negative burglary is not a viable concept the linear regression, although still very significant, is in this case biased by an overlarge set of zero burglary entries, leading to a lower regression slope parameter for both unique and repeat cases than might otherwise be expected.

It is clear that the values of the regression slopes have stabilised once EDs with no burglaries have been omitted, as seen in the unique and repeat sets in Table 7-7. These two sets show great consistency and separability of slope value. The slope of the regression for repeat victimisation occurrences is about four times steeper than that for unique events (Table 7-7 and Figure 7-11). The increase in deprivation index as unique burglary numbers increase is definite but much less rapid. The slope values in Table 7-7 would indicate that although unique occurrences of burglary are widespread throughout Trent Division, there is a definite but limited relationship to increasing social deprivation. Increasing numbers of repeat victimisation cases, on the other hand, are clearly indicative of rapidly increasing social deprivation. Area deprivation values increase slowly with a rise in the unique event rate, but increase rapidly with rising repeat victimisation.

The intercept values in Table 7-7 are very similar especially for the more "reliable" unique and repeat ED trials. The unique trial contains the complete data set for unique and repeat burglary locations without the bias problem caused in the valid ED set and can therefore be used to demonstrate the rate of increase as EDs become more socially deprived. The extent of the data cloud and associated linear regression for each data set is plotted in Figure 7-11.

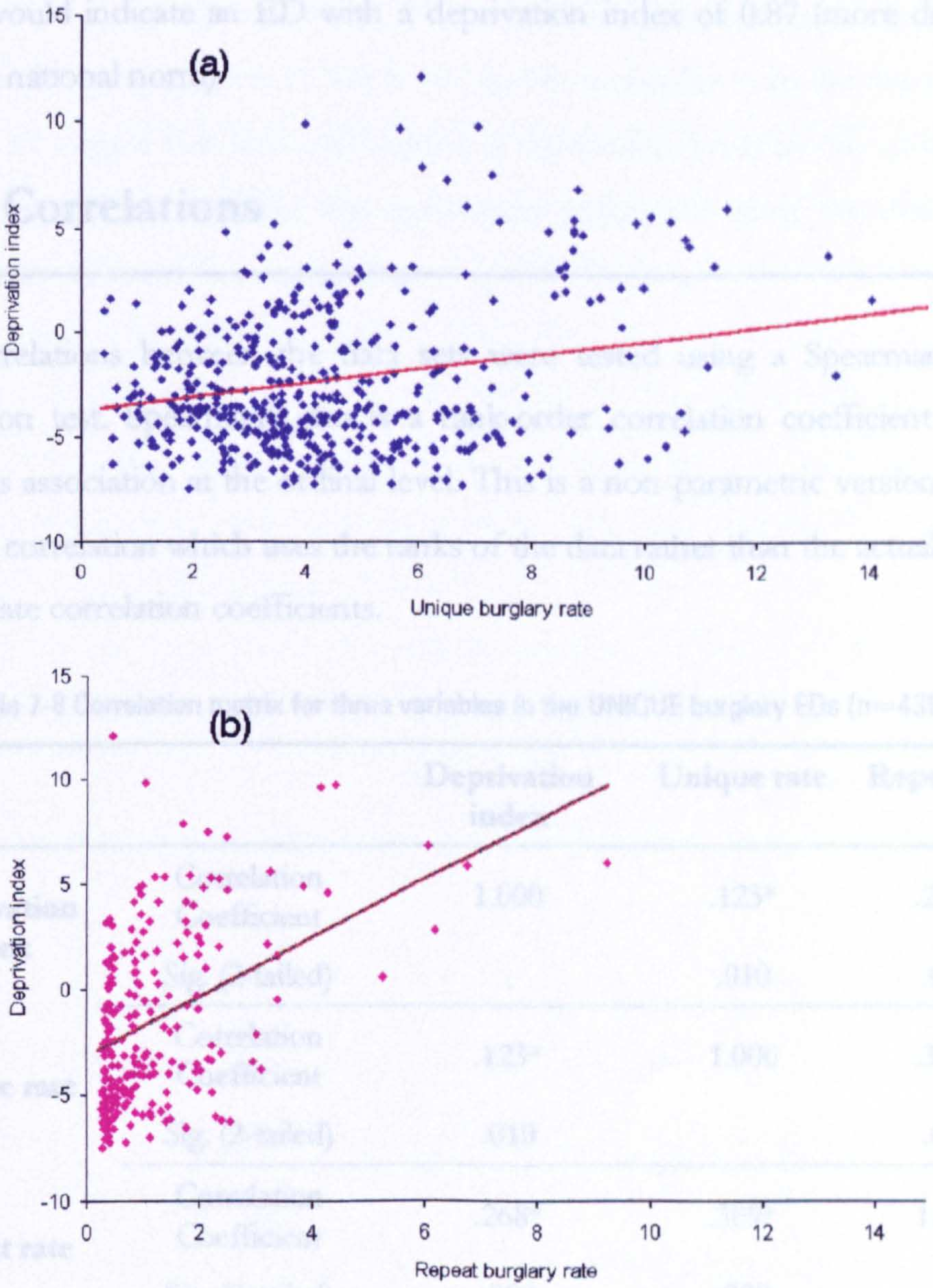


Figure 7-11 Data clouds and linear regression lines for unique (a) and repeat (b) burglary rates.

As the burglary rate increases the repeat Vicinity values show a more rapid increase in deprivation index value, indicating that households which are vulnerable to repeat victimisation tend to be in more deprived areas.

Both lines intercept the deprivation (y) axis within less than one deprivation index unit of each other for zero burglaries. This indicates that in all cases more affluent (negatively deprived) EDs are those tending strongly to have low burglary rates. At a burglary rate of zero the regression lines are less than one unit apart at a value of about -3.5. Once a 3% unique rate of burglary (found at the lower end of the unique data cloud) has been reached the expected deprivation score for a given ED will be -2.88 (still affluent), whereas the same 3% rate for repeat victimisation (towards the high extreme of the repeat data cloud) would indicate an ED with a deprivation index of 0.87 (more deprived than the national norm).

7.2.5. Correlations

The correlations between the data sets were tested using a Spearman's rho correlation test. Spearman's rho is a rank-order correlation coefficient which measures association at the ordinal level. This is a non-parametric version of the Pearson correlation which uses the ranks of the data rather than the actual values to calculate correlation coefficients.

Table 7-8 Correlation matrix for three variables in the UNIQUE burglary EDs (n=439).

		Deprivation index	Unique rate	Repeat rate
Deprivation index	Correlation Coefficient	1.000	.123*	.268*
	Sig. (2-tailed)	.	.010	.000
Unique rate	Correlation Coefficient	.123*	1.000	.309*
	Sig. (2-tailed)	.010	.	.000
Repeat rate	Correlation Coefficient	.268*	.309*	1.000
	Sig. (2-tailed)	.000	.000	.

* *Correlation is significant at the .01 level (2-tailed).*

Table 7-8 displays correlation coefficients and 2-tailed significance values. The significance level is the probability of obtaining results as extreme as the one observed when there is no correlation. If the correlation displays a significance level of less than 0.05 then the correlation is considered to be significant and the two variables are linearly related. If the correlation has a level of significance which is relatively large (for example, 0.50) then the correlation is not significant and the two variables are not linearly related.

Table 7-8 shows a significant linear relationship between all of the three variables (deprivation index, the unique burglary rate, and the repeat burglary rate) for all of the 439 enumeration districts containing a burglary. Each of the variables demonstrates a positive correlation. The results from this test can be interpreted alongside the results shown in Table 7-7, and these results from the original data appears to suggest that although there is a correlation between the unique and repeat rates of burglary and the deprivation index, the linear regressions are significantly different from each other.

7.3. DISCUSSION

By focussing on the locations of incidents this study has determined that the social deprivation in the vicinity of unique burglary sites and repeat burglary sites is significantly different. There is a significant difference in the means of the distributions, and also in the frequency distributions. The joint pattern of household and unique burglary locations compared against the deprivation index is very clear and shows a significant but slight increase in the rise of burglary risk as the level of deprivation increases (Figure 7-11a). The sites where repeat victimisation has occurred show a marked difference (Figure 7-11b). As social deprivation increases, the relative possibility of being a victim of repeat burglaries increases dramatically. The question arises as to why the occurrences of burglary repeat victimisation are concentrated in the regions of greater social deprivation.

There are a number of possible hypotheses to explain this. The availability (or not) of crime prevention resources is one such explanation for this phenomenon. It has been recognised for a number of years that crime prevention programs can help to reduce the occurrence of burglaries and repeat incidents (Kennedy and Veitch, 1997; Spelman, 1995), though the availability of crime prevention programs is a finite resource. If more deprived areas have higher crime rates (as is often the case), then the available money and assets to conduct preventative work after an initial burglary may be spread between many more locations, or may not be available at all. Local authorities in deprived areas have a limited amount of capital and crime prevention is only one of a number of necessary expenditures. In more affluent areas, the lack of resources centrally may prompt household owners themselves to invest in crime prevention measures after an initial burglary, an option not financially available to more indigent residents. This understandable self-interest in protecting the home can be seen in any affluent suburban street by comparing the number of houses with burglary alarm boxes to the number of houses with alarms in a poorer inner-city area.

Style of policing could also be a factor in the results seen in this chapter. Recent work has highlighted the effect that policing style can have on the crime level in an area. There is evidence to suggest that police response to incidents and general policing style is affected by the level of crime in the area (Klinger, 1997). Officers who police busy areas tend to be diverted to the more serious calls and may ignore minor legal infringements which would otherwise be dealt with more rigorously in a less crime-ridden district. Constant attendance at serious incidents leads the officers in busy areas to conclude that they are in a crime-ridden area and may lead to a belief that a high burglary rate, and a high level of repeat victimisation, is an inevitability in a deprived area. This feeling of crime as inevitable can be passed on to the local population who may likewise feel that the additional cost and effort of crime prevention measures may be worthless.

In more affluent areas, police officers may perceive a lower crime rate and consequently have more time to perform routine patrolling. They may even be able to target actively their patrol routes to keep an eye on recently burgled premises. The author has some personal experience of this. Prior to entering academia the author worked for 10 years in both Tower Hamlets and Central London as a police officer. Central London exhibited a low burglary rate and each incident was actively pursued by the local force with follow-up visits, crime prevention advice and increased patrolling. During the late 1980s and early 1990s Tower Hamlets was the most deprived district in the UK (Environment, 1995) and the burglaries rate was huge. The sole crime prevention officer only attended those locations where the reporting officer felt any benefit would be achieved, and the sheer volume of criminal activity in the area meant that additional patrols were usually only ever increased for violent and drug-gang related incidents. Burglary was seen as inevitable.

7.4. CONCLUSION

This chapter aimed to examine the differences between locations burgled once, and repeat victimisation burglaries, in the context of social deprivation. It has been shown that there is a significant difference between the social deprivation index of areas in the immediate vicinity of once-only burgled premises and the deprivation index in the vicinity of repeat victimisation sites. Locations which were subjected to repeated attacks during the two year study were in significantly more deprived areas than the unique burglary premises. There was a marginal difference in the measure of deprivation index at the burglary sites and in the immediate vicinity of the site, and to prevent the effect of peninsularisation the study has stressed the benefit of employing a *Vicinity* approach. This type of analysis uses an areally weighted average to reduce the effect and is designed to produce a reasonably continuous distribution of the target variable. As a result of the tests a radius of 100 metres is suggested for urban areas.

A number of hypotheses are suggested for the cause of increased repeat victimisation in deprived areas, and these include the lack of crime prevention resources in deprived areas, the possibility of more affluent residents being able to afford to take responsibility for preventative measures, and the difference in policing styles between affluent and deprived areas. Any one or more of these reasons may be the cause, or as is more likely, the cause of this phenomenon is a combination of many factors.

8. Hotspot analysis

The thesis now proceeds to investigate the area of hotspot analysis. This is a growing field of interest within crime analysis and there are a number of reasons why hotspot analysis is important to policing. These reasons include the value in accurately aggregating both high volume and repeated locations of crime to make the information more comprehensible to patrol officers, and the ability to identify areas for crime prevention activity.

Before this type of analysis can become useful, the accurate identification of hotspots is essential and this chapter examines the available software and algorithms before presenting a new method for the identification of crime hotspots.

8.1. INTRODUCTION

There are a number of areas where hotspot analysis can be of use to the police service. The ability to aggregate high volume crime and present patrolling officers with a simplified map emphasising the important areas to police is useful in preventing 'information overload'. The use of simple hotspot maps also allows the police to share sensitive crime location information with outside parties such as the local council (who are interested in crime prevention) without compromising the confidentiality of the data and revealing individual locations. This chapter evaluates the current software and algorithms for hotspot mapping, and presents a new method based on a local statistical algorithm.

Recent years have seen an increase in police workload due to a more diverse role (public order, community education, traffic management) and an expansion in administration (for example the additional responsibilities under the Police and Criminal Evidence Act 1984). This has not been matched by an increase in police numbers (Morgan and Newburn, 1997). These factors have combined to mean that the available time an officer has for routine patrolling has been diminished. It is vital therefore that the best use is made of the time available and that the officer is familiar with the most likely sites on his beat for criminal activity; usually the areas of highest incident concentration. It is commonly the duty of a local intelligence officer or management information unit to disseminate this type of information to operational officers. Burglary and motor vehicle crime are often highlighted as they tend to be both national and local priorities due to their high volume nature, yet a long list of crime locations and times on an office wall is unlikely to impart much wisdom to even the keenest probationer. This is common practice in Nottinghamshire, as can be seen from Figure 8-1 which shows a whiteboard used at Clifton police station (photograph taken February 1998 and shown at this resolution to make the sensitive information deliberately illegible) to update the next shift with the most recent incidents. It is the only visual medium used at the station for information dissemination, with all other briefings being in a written form or read to the oncoming shift by the duty Sergeant. These briefings usually follow a standard format where the new shift are informed of: all burglaries and motor vehicle crimes occurring since the shift were last on duty, details of any major incidents, suspicious people or vehicles to identify, properties to pay particular attention to, and any other business. Attempting to extract an accurate picture of crime distribution from this mass of information is extremely difficult.

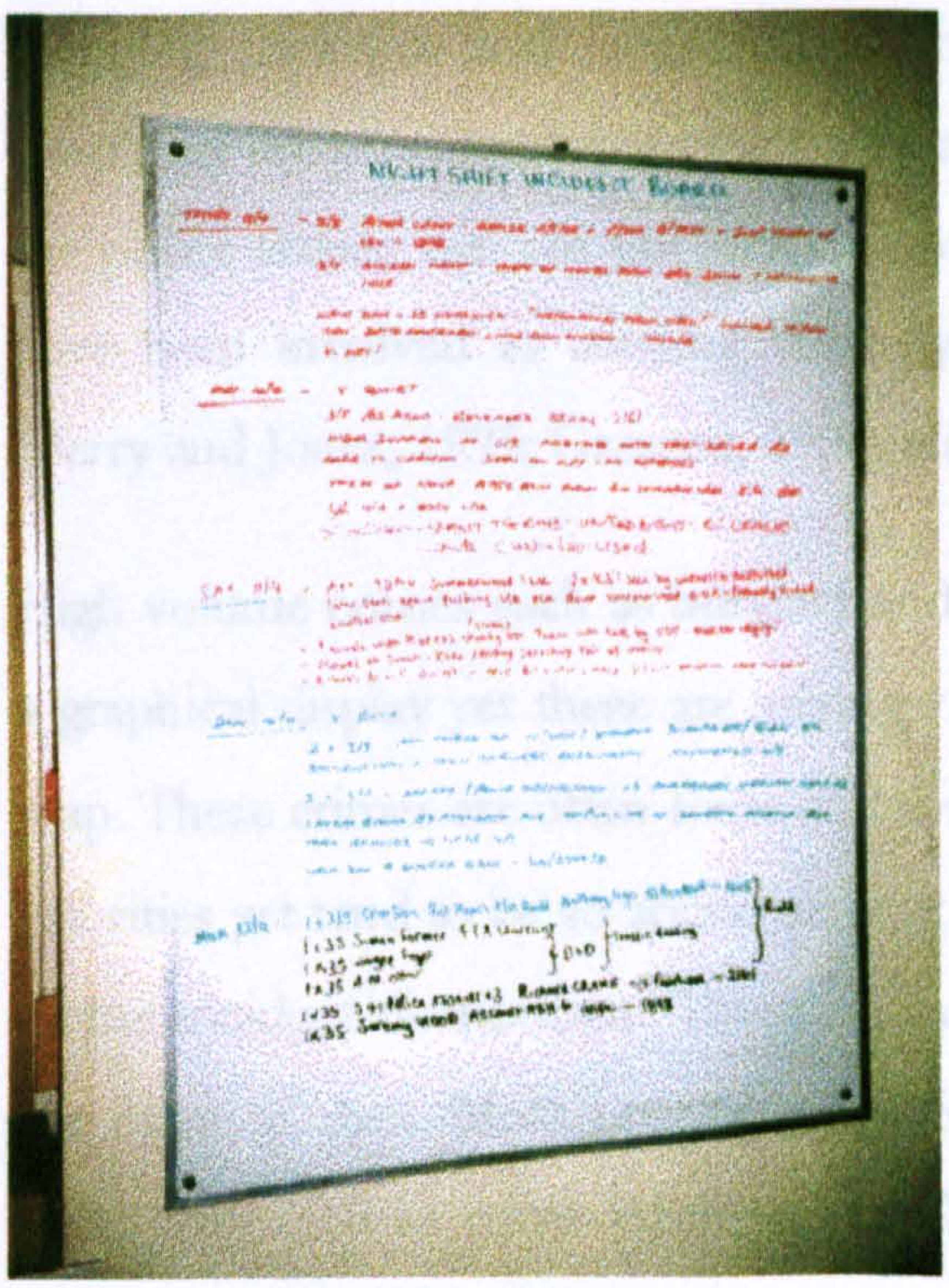


Figure 8-1 Whiteboard used for information dissemination at Clifton Police Station (Feb 98)

If you walk into most police stations in Great Britain today you are likely to see maps of one form or another. It is obviously important for officers to know where they are, and so from the moment they arrive at a station they are usually given a street map of the station area with which they familiarise themselves. This fits in with the current mode of policing in Britain which is geographical in nature, from the county constabulary down to the single beat. The knowledge of where past crimes have been committed is an important part of identifying higher risk areas in the hope of preventing future incidents and so a map is an excellent medium for this information, assuming that police officers are familiar with the map in the first place. The realisation that a map is an ideal visual medium for crime distribution is not a new one.

Traditional crime mapping has involved sticking pins into a map to show the location of crime incidents. This has a number of positive attributes for the police: it is simple, accurate, easily understood, not prone to error (unless the pins fall out) and anyone can do the pinning. There are a number of problems with this. When the map is full there is often too much information, and when you take away the pins and start again (for example at the beginning of a month) there is too little information. Too many different crime types present a mass of colours and confuse the viewer, and numerous crimes at one location also present difficulties. It is not an ideal solution.

In response to this a number of forces are investigating the use of digital mapping systems to display crime information. This is now possible as many forces include geographical data such as grid references with their recorded crime and incident information. The use of georeferenced data in digital mapping packages and Geographical Information Systems (GIS) has vastly simplified the process of mapping crime incident data. In this manner the user at the police station can see the locations of every incident. A number of authors have been involved in creating these types of crime-based mapping systems (Berry and Jones, 1995; Grescoe, 1996; Hirschfield *et al.*, 1995; Project, 1997).

High volume crimes such as burglary and motor vehicle crime can often swamp a graphical display yet these are amongst the most important types of crime to map. These crimes are often force and divisional crime prevention and detection priorities yet tend to be so abundant that it is difficult for officers to get an idea of the incident distribution. Thematic mapping is one solution to coping with the mass of data. Most proprietary GIS programs include the ability to map point data such as crime information. This can cause different descriptions of the data due to both the aggregation method and the modifiable areal unit problem (MAUP), which is discussed later in this chapter in section 8.1.2.

8.1.1. Thematic mapping

There are a number of different ways of creating thematic maps by aggregating data. The simplest technique is to count events occurring in a predefined area where the number of incidents within a polygon (such an a police beat) are counted. For example, programs such as MapInfo (a commercial GIS and mapping package) offer a number of pre-programmed techniques for aggregating area data. Thematic classifications include different display variations of the count per polygon calculation. These include; equal ranges, quantile, standard deviation and natural break method. Each of these methods can produce different cartographic outputs. This can be seen in Figure 8-2 which shows the same data for West Bridgford using four of the automatic thematic map options available in MapInfo. All recorded crime for West Bridgford for the period April 1995 to April 1997 is shown aggregated by beat (a), and by 500 metre squares (b, c and d). As can be seen from these four images, different choices of aggregation method in the creation of a thematic map can affect the visual outcome.

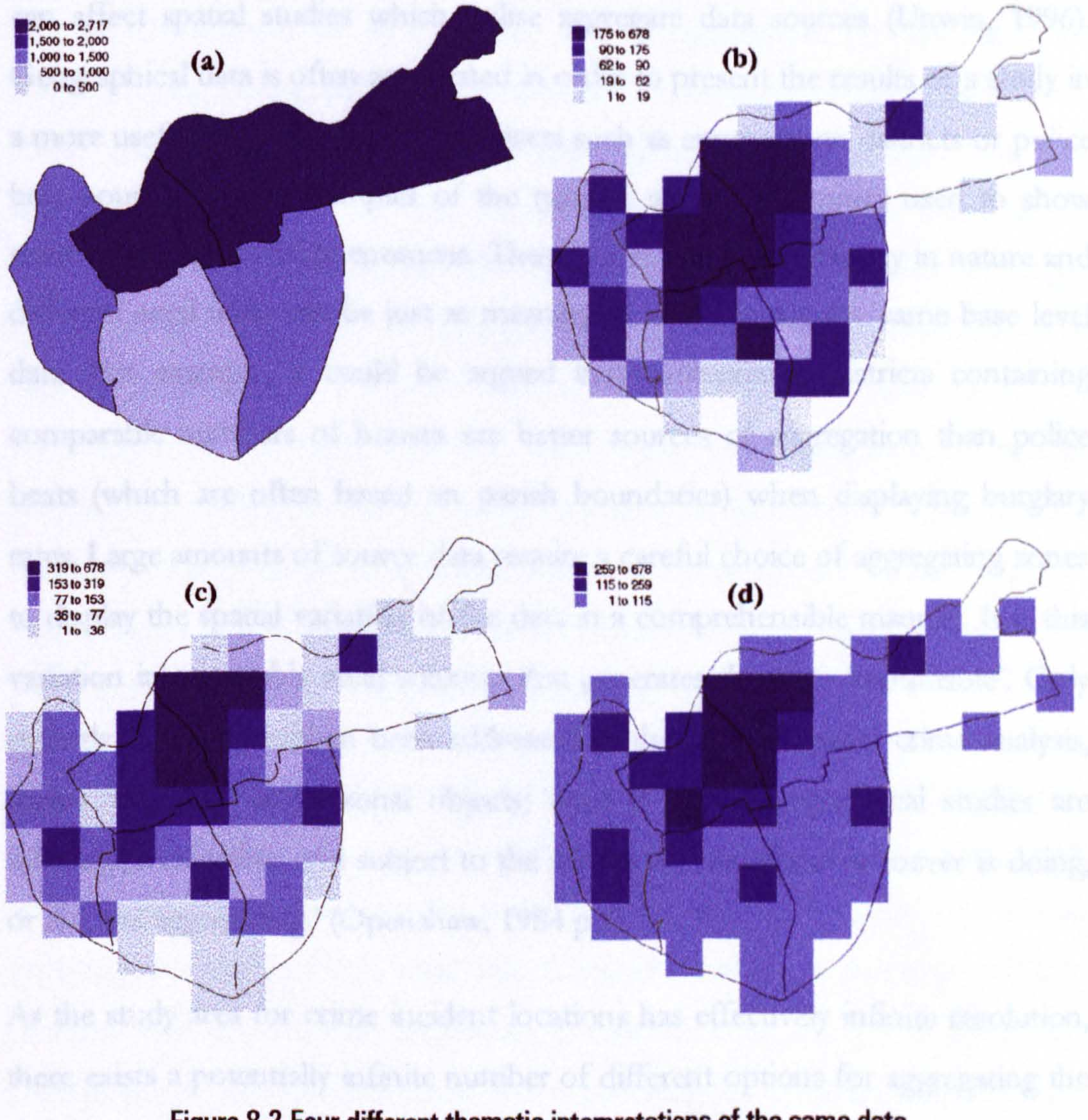


Figure 8-2 Four different thematic interpretations of the same data.

Figure 8-2 (a) shows all recorded crime for West Bridgford from April 1995 to April 1997, aggregated by police beat. (b) shows the same data using the equal count method, (c) the natural break method, and (d) the categories broken down into standard deviations.

The differences between (a) and the grid square images (b, c & d) is a simple example of the Modifiable Areal Unit Problem, which is explained in greater detail in the next section. With image (a), a viewer might be fooled into thinking that a high level of crime extends right across the top right police beat. However reference to the other images show that the high level of crime is concentrated to the extreme West of the police beat, and that the Eastern section of the beat is almost crime free.

8.1.2. The Modifiable Areal Unit Problem

The modifiable areal unit problem (MAUP) is a potential source of error that can affect spatial studies which utilise aggregate data sources (Unwin, 1996). Geographical data is often aggregated in order to present the results of a study in a more useful context, and spatial objects such as enumeration districts or police beat boundaries are examples of the type of aggregating zones used to show results of some spatial phenomena. These zones are often arbitrary in nature and different areal units can be just as meaningful in displaying the same base level data. For example, it could be argued that enumeration districts containing comparable numbers of houses are better sources of aggregation than police beats (which are often based on parish boundaries) when displaying burglary rates. Large amounts of source data require a careful choice of aggregating zones to display the spatial variation of the data in a comprehensible manner. It is this variation in acceptable areal solution that generates the term 'modifiable'. Only recently has this problem been addressed in the area of spatial crime analysis, where 'the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating.' (Openshaw, 1984 p.3).

As the study area for crime incident locations has effectively infinite resolution, there exists a potentially infinite number of different options for aggregating the data. Numerous administrative boundaries exists, such as enumeration districts, wards, counties, health authority areas, etc. Within modern GIS, it is an elementary task to automatically generate a huge variety of non-overlapping boundaries. Regular, often square, grids are common, though polygons have been used in other studies of crime distribution (Hirschfield *et al.*, 1997). The number of different combinations of areal unit available to aggregate data is staggering. Openshaw (1984) calculated that if one was to attempt to aggregate 1,000 objects into 20 groups, you would be faced with 10^{1260} different solution combinations. Although there are a large number of different spatial objects and ways in which a large geographical area can be sub-divided, the choices of areal units tend to be dominated by what is available rather than what is best. Police crime data is often mapped to beats, even when the information is passed to outside agencies such as neighbourhood watches or local councils who might benefit from more relevant boundary structures.

The MAUP consists of both a scale and an aggregation problem, and the concept of the ecological fallacy should also be considered (Bailey and Gatrell, 1995). The scale problem is relatively well known. It is the variation which can occur when data from one scale of areal units is aggregated into more or less areal units. For example, much of the variation in enumeration districts changes or is lost when the data is aggregated to the ward or county level. The aggregation problem is less well known and becomes apparent when faced with the variety of different possible areal units for aggregation. Although geographical studies tend towards aggregating units which have a geographical boundary, it is possible to aggregate spatial units which are spatially distinct. Aggregating neighbours improves the problem to a small degree but does not get round the quantity of variation in possibilities which remains.

The ecological fallacy problem occurs when inference is made about individuals based on the results of a study which has aggregated data into areal units. Whether this problem may or may not exist depends on the data and the choice of aggregation method and areal unit. A simple example of this problem is to consider two hypothetical enumeration districts with the same population demographics. Every economically active worker in ED1 has an annual income of £10,000 while every worker in ED2 has an income of £50,000. If these two enumeration districts were aggregated into one ward, then the mean income of the economically active occupants of the ward is £30,000. It would however be an ecological fallacy to suggest from this that most workers in the ward had an income of about £30,000.

SOLUTIONS?

> *It is not thought likely that a general solution can be found that will allow existing methods to be used as if the MAUP did not exist. The problem is far too complex, it is difficult to investigate by analytical means, and its inherent geographical nature makes it unlikely that a statistical solution will emerge or if it does that it will suffice. (p.31)*

> *It is a geographical fact of life that the results of spatial study will* always *depend on the areal units that are being studied. (Openshaw, 1984, p.37)*

A temporary local solution is possible where an agency, such as Nottinghamshire Constabulary, agrees on a standard spatial unit on which to base every geographical analysis. This would only be of benefit to the police, as outside agencies would need to agree either to copy the chosen police boundary or else risk finding police data incompatible with their own 'internal' boundary solution. It would also be a short term measure as boundary changes are needed from time to time to reflect the changing population pattern and the adaptation in techniques and methods of policing. Such a boundary change took place within Nottinghamshire Constabulary during the course of this thesis, and is detailed in chapter 3 (Data sources and software).

Further work on the MAUP is necessary, however at present for the purposes of this thesis the most promising empirical solution is to examine geographical phenomena at a variety of scales and spatial units. In this way, spatial patterns which truly exist should be manifest at a variety of spatial resolutions.

8.2. Point Pattern Analysis

The development of computing has seen the introduction of spatial analysis techniques which were previously impossible without the ability to perform a large number of iterations. The following discussion of spatial analysis is summarised from Gatrell *et al.* (1996).

Spatial point pattern analysis focuses on the concept that a set of events have occurred at given locations within a study area. An analysis takes place within the region in an attempt to determine the locations of any possible clusters. This study region is often rectangular in shape as this design is easy to program, and the design of computers tend to favour matrix calculations. Whatever the shape of the study region constructed, the possibility of edge effects in the analysis must be considered.

8.2.1. Edge effects

The problem of edge effects is demonstrated in Figure 8-3, where the top left corner of a study area is shown. Test circle A is able to select events from within 100% of its circumference, however test circle B is centred over the extreme top left corner of the study region. This circle is restricted to including all possible events within 25% of its area, as no incidents fall outside the study region. This can be countered by leaving a guard area between the perimeter of the study area and a resulting sub-region which becomes the effective new study

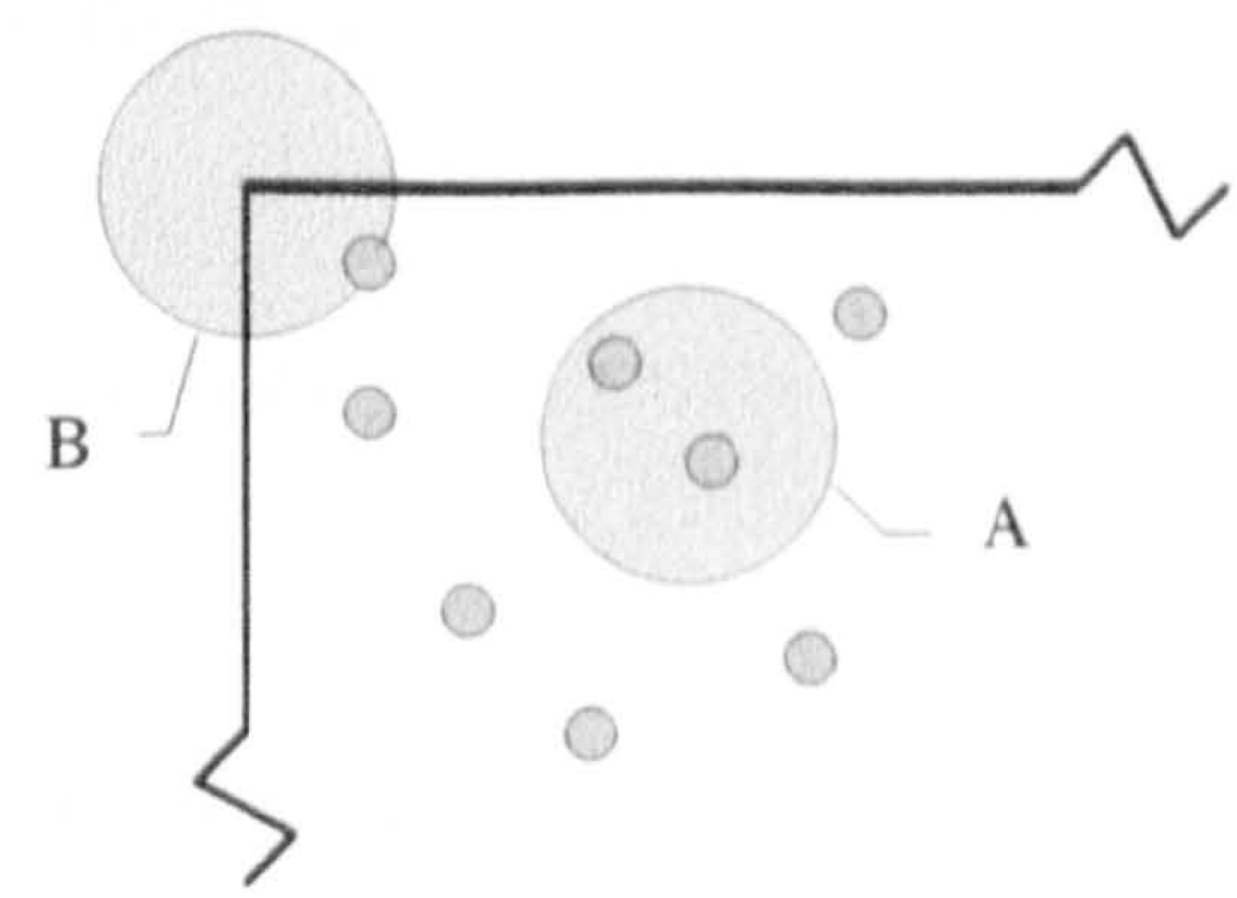


Figure 8-3 The top left corner of an example study region is shown, with eight 'events' depicted.

Test circle A is able to select events from the full range of its area, however test circle B is affected by edge effects as it is only able to select events from within 25% of its area.

area. Alternatively, the analysis technique can be programmed to take account of edge effects. If this were done, the test circle B in Figure 8-3 would allow for the fact that it can only 'see' a quarter of the study region. While the example of a circle over a corner is easy to calculate (25%) it becomes much more complex to calculate the area of the circle within the test area when you either move in from the edge, or have an irregularly shaped study region. Under these circumstances it is easier to either reduce the study area or increase the range of the event selection beyond the study region.

The null hypothesis in this type of analysis is the contention that events are distributed randomly across the study area according to a uniform probability distribution. This theoretical concept of complete spatial randomness is unlikely to be observed in crime data and therefore any departure from this concept will display either regularity or clustering. The next stage is to examine methods of calculating the measure of regularity or clustering.

A simple spatial analysis might place a grid over the study area and count the number of crimes in each square. This is a standard point counting technique which is often done in geographical applications. The problems with this are described in an earlier section and shown in Figure 8-2 on page 215.

However instead of doing this, a more advanced technique is to place a moving window, such as a circle, over the study area, and this is called the 'moving window' technique.

8.2.2. Moving window technique

The 'moving window' technique employs a moveable sub-region (usually a circle) over the entire study area to measure dependence in subsets of the study area and is particularly suited to hotspot detection (Bailey and Gatrell, 1995). A two-dimensional grid lattice which covers the entire study area with a rectangular grid of intersecting lines is defined and at each grid intersection circles are placed over the study area. Points falling within the circle are retrieved from the data to compute a spatial pattern test statistic. This process is seen in Figure 8-4, where

(a) shows a fictitious street network with the location of burgled premises shown as red dots. The moving window analysis approach is to superimpose a grid over the whole study area (b) and then at each grid intersection overlay a circle (or circles) of a predetermined radius (c) – commonly referred to as a 'bandwidth'. The crimes (points) falling within the circle are extracted and some form of algorithm is applied to extract a value. This could be a simple count of the number of crimes in the circle, a density calculation based on the number of incidents and the area of the circle, or some more complex algorithm such as a kernel estimation, which will be discussed in a later section.

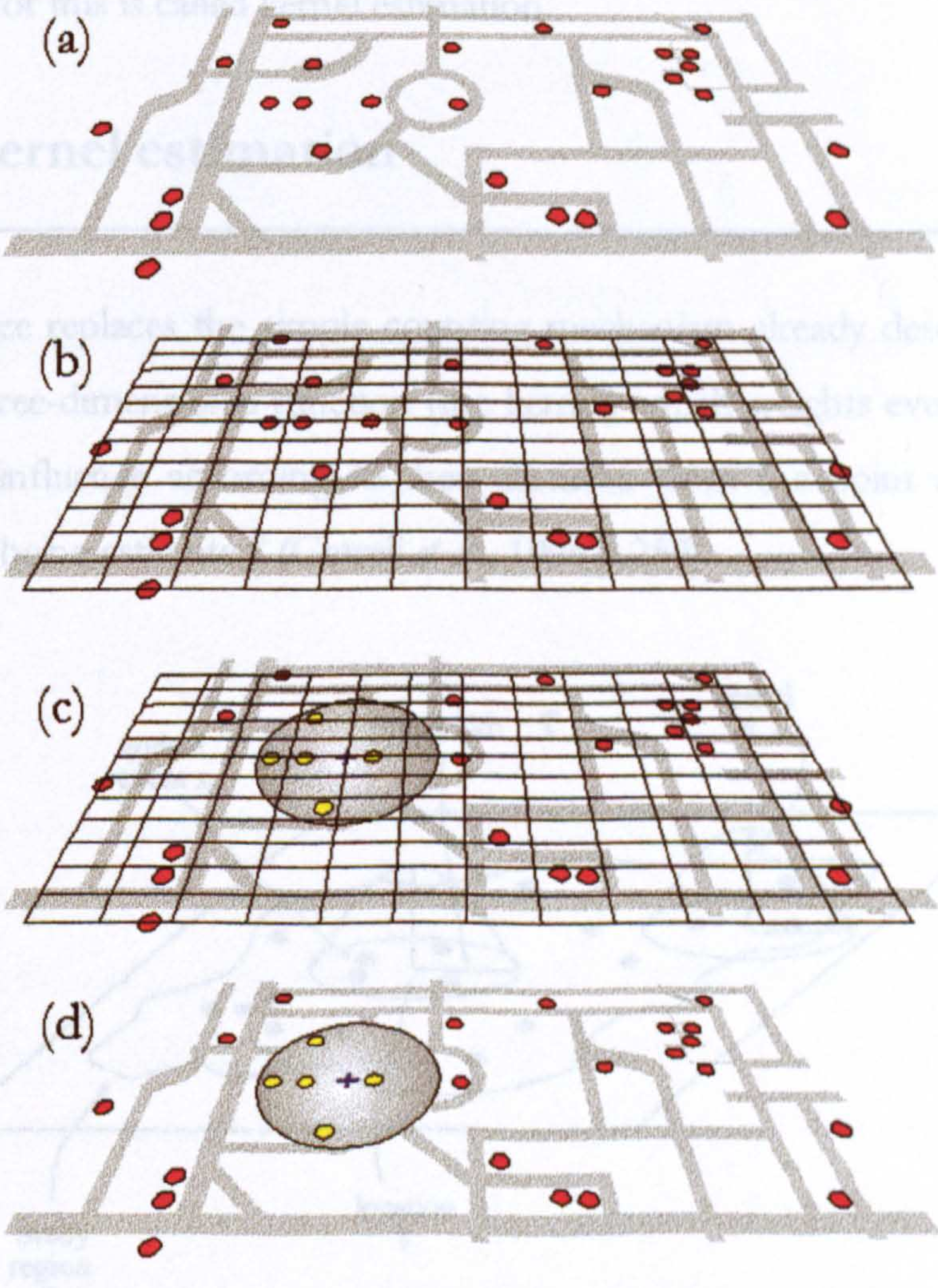


Figure 8-4 Moving window spatial analysis process.

This technique produces a more spatially smooth estimate of the variation than can be obtained with a fixed quadrat system such as the choropleth mapping technique. In addition to this the use of a moving window where windows overlap can help defeat much of the MAUP (Bailey and Gatrell, 1995;

Openshaw, 1984; Unwin, 1996). It is also a local method in a crime analysis environment where global methods are difficult to justify. Given the 'distance-to-travel' limitations of many offenders (Bradbury, 1981; Brown, 1982) local methods are more desirable than global methods which relate crime locations right across the study area and can interpolate values into areas void of criminal activity.

In each of the density calculations however, no account is taken of the relative location of events within the search window. Density calculations can be a useful tool, though an improvement is the ability to calculate an intensity measure. One technique for this is called kernel estimation.

8.2.3. Kernel estimation

This practice replaces the simple counting mechanism already described with a 'moving three-dimensional function (the kernel) which weights events within its sphere of influence according to their distance from the point at which the intensity is being estimated' (Gatrell *et al.*, 1996 p.259).

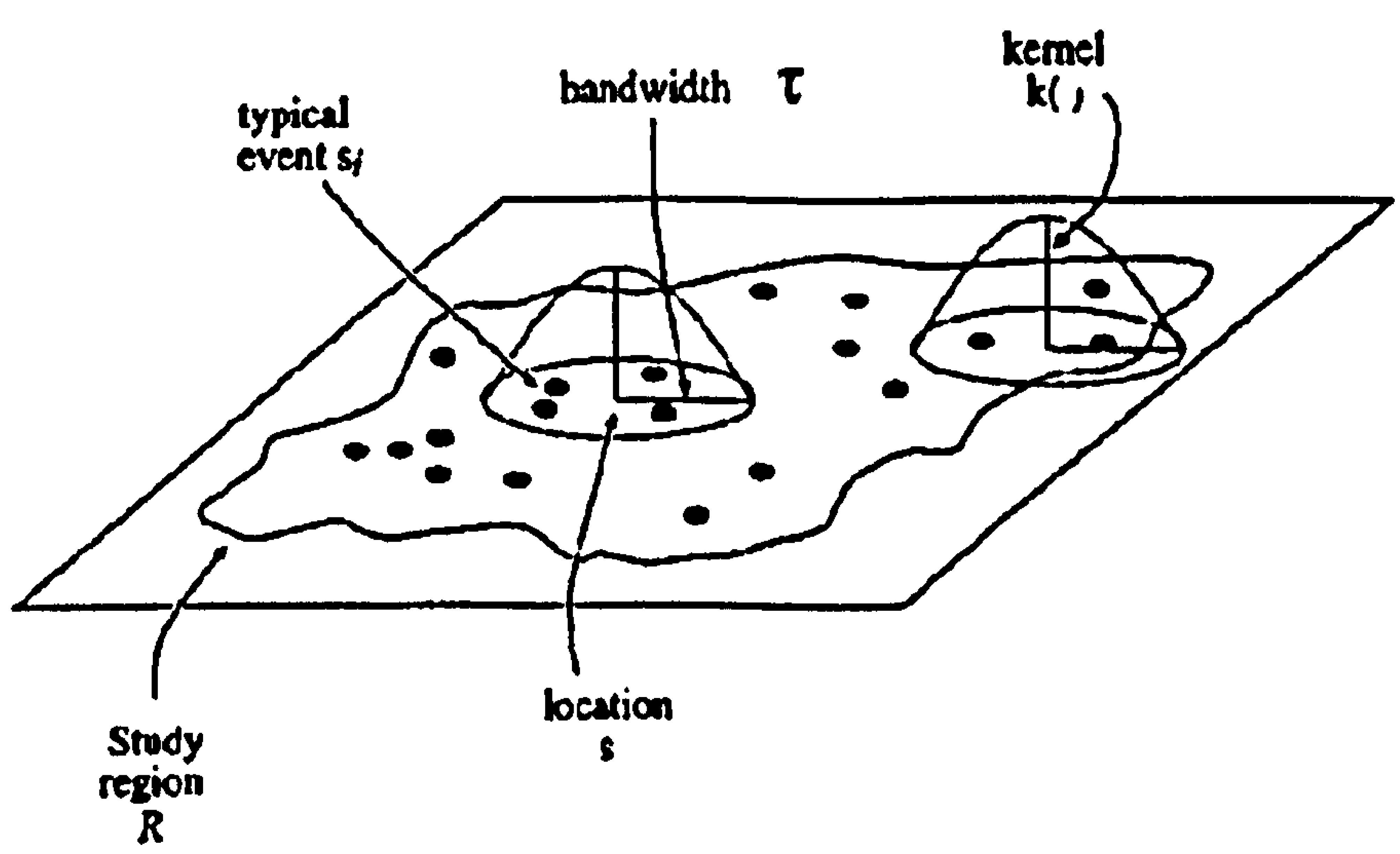


Figure 8-5 Kernel estimation of a point pattern.

Source: Gatrell et al., *1996.*

The three-dimensional function scans within the search circle and not only detects points within the search region but measures their influence and calculates their contribution to the intensity of the search relative to their proximity to the centre of the search circle. The closer an event is to the centre of the circle, the greater its contribution to the intensity reading. Options for displaying the results include the use of a contouring program to show the results as an intensity surface, or to select the locations and circle sizes with the highest readings. Figure 8-5 shows this process where *k* represents any kernel weighting function which decays away from the centre or the search location at location s. The bandwidth τ can be adjusted to vary the range of influence of the search circle. Figure 8-6 shows the equivalent figure for the sequence in Figure 8-4 and completes the picture.

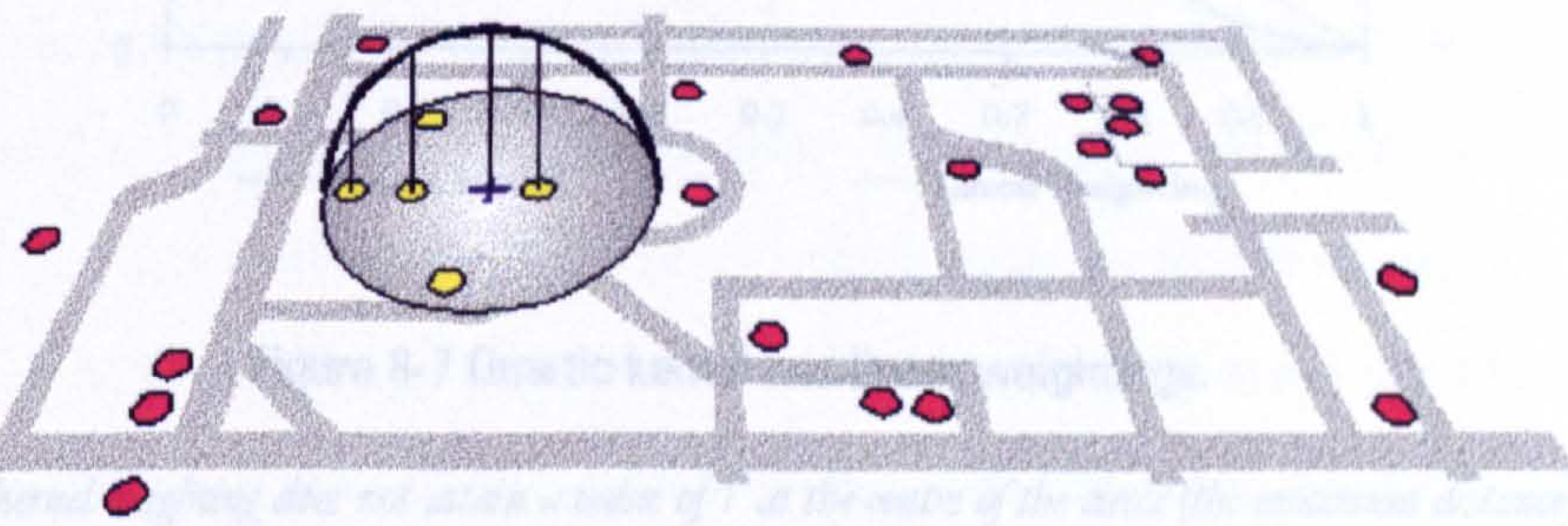


Figure 8-6 An inverse distance weighted function (kernel) applied to crime locations.

$$\hat{\lambda}_{\tau}(s) = \sum_{d \le \tau} \frac{3}{\pi \tau^2}\left(1 - \frac{d^2}{\tau^2}\right)^2 \qquad \text{Equation 8-1}$$

The choice of suitable algorithm for the search parameter has, according to Gatrell *et al.* (1996), little bearing on the resulting intensity estimate. Later in this chapter a local surface generation program written by the author called SPAM (Spatial Pattern Analysis Machine) will be described. The algorithm used by SPAM is a quartic kernel algorithm from the Gatrell *et al.* (1996) paper as shown in Equation 8-1, where *s* represents the centre of the search circle, τ the bandwidth and *d* is the distance of each point (*i*) within the bandwidth (τ) from the centre of the search area (*s*). The calculation of the intensity $\lambda_\tau(s)$ is therefore the summation of the intensity of those values which have a smaller distance from s than d, where the intensity is weighted, with values at the centre weighted by $3/\pi\tau^2$, dropping smoothly to a value of zero at the maximum distance τ. This has been suggested as a reasonable choice and useful in a variety of typical applications (Bailey and Gatrell, 1995 p.85). The function shows a weighting that favours points close to the centre of the circle and decays smoothly towards the perimeter.

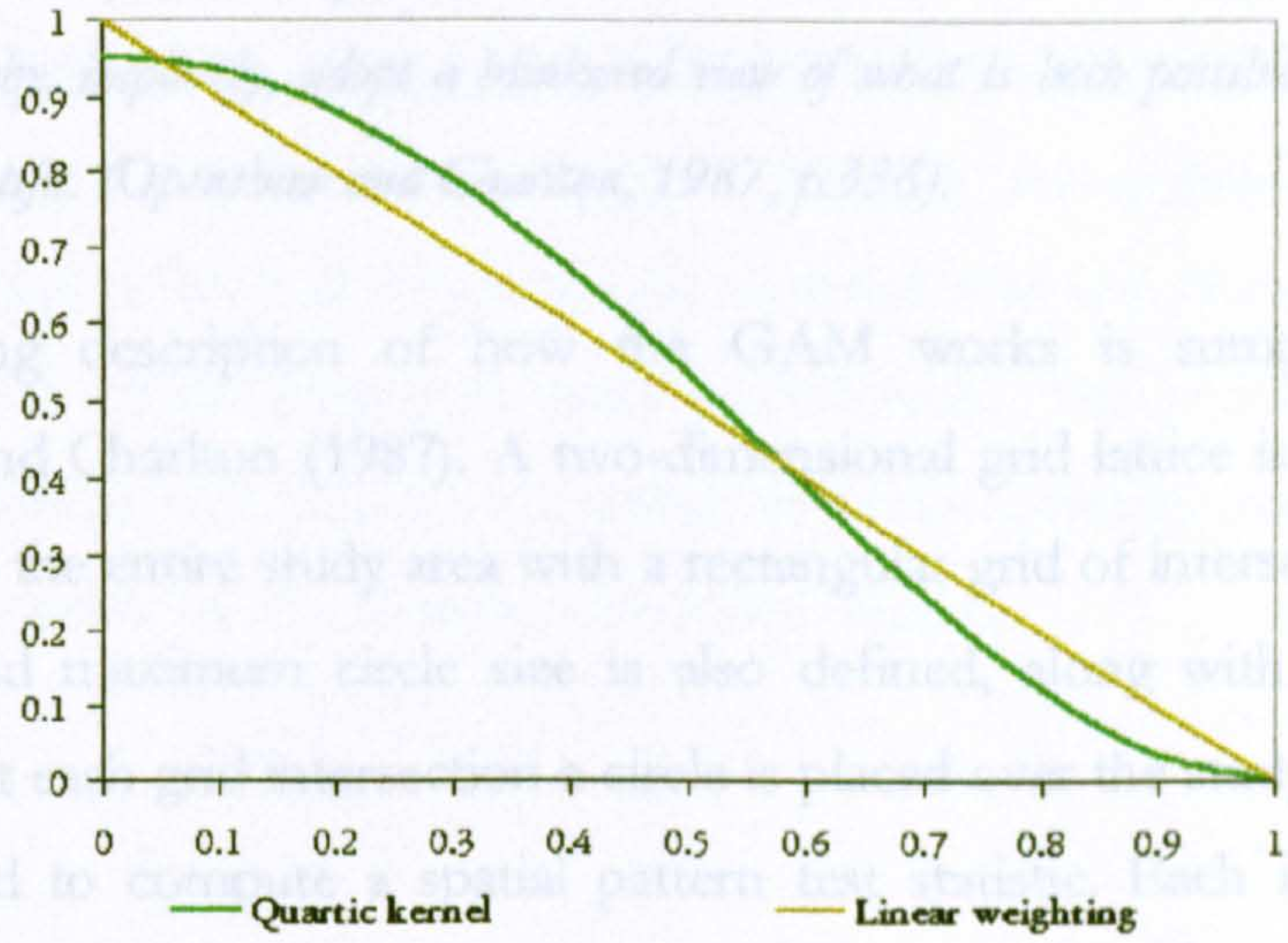


Figure 8-7 Quartic kernel and linear weightings.

The quartic kernel weighting does not attain a value of 1 at the centre of the circle (the minimum distance) as the weighting here is $3/\pi$.

Much of the work into these spatial analysis methods has been driven by the field of medical epidemiology and the desire to detect clusters in rare diseases (Besag and Newell, 1991; Gatrell *et al.*, 1996). One of the more influential attempts have been the Geographical Analysis Machines designed by Stan Openshaw and his colleagues at the University of Leeds (Openshaw and Charlton, 1987; Openshaw *et al.*, 1998).

8.2.4. GAM Mark I

The GAM Mark I was designed to be an automated process by which point data, fed into a computer as co-ordinate values, could be searched automatically by various algorithms to detect clusters (Besag and Newell, 1991; Openshaw and Charlton, 1987). GAM was introduced originally to investigate the number of leukaemia cases around power stations in the north-east of England in the early 1980's. The new technique was developed because the authors felt that with;

> *the development of GIS technology and the vast growth in spatially-referenced data sets, the available tools are no longer adequate for the tasks of data processing that they now face. There is also the tendency to limit the spatial analytical tool-kit to established statistical methods and thereby, implicitly, adopt a blinkered view of what is both possible and scientific. (Openshaw and Charlton, 1987, p.336).*

The following description of how the GAM works is summarised from Openshaw and Charlton (1987). A two-dimensional grid lattice is defined such that it covers the entire study area with a rectangular grid of intersecting lines. A minimum and maximum circle size is also defined, along with a radial size increment. At each grid intersection a circle is placed over the study area and the data retrieved to compute a spatial pattern test statistic. Each intersection is visited and the test statistic for the current circle size is calculated for the whole grid. The GAM then changed to the next circle size and repeated the whole test. It continues in this way until every grid intersection has been visited by each of the chosen circle sizes. For the particular epidemiological study for which the GAM was designed, the test statistic has to consider the background population density. The GAM generates a number of randomly generated values, based on the underlying population-at-risk density, and a Monte Carlo procedure for significance testing is employed to detect clusters to a significance level of 0.01. Among the important features of GAM is the lack of a need for a prior hypothesis. The program allows for the trawling of data for interesting features, removing the need for the operator to consider a possible line of enquiry and then test for it.

Though the GAM has been criticised for some of the statistical assumptions (Besag and Newell, 1991), it still remains one of the main innovations in recent spatial analysis.

A recent attempt to implement the GAM for crime analysis has been made by the research team at the University of Liverpool. They reprogrammed the GAM as part of a larger project into creating a *profiler* to investigate crime and urban deprivation (Hirschfield *et al.*, 1997). The profiler is a software package that combines a number of databases into an analytical system. In the late 1980's the GAM, while conceptually simple, was computationally intensive and runs could take days. It was originally run on an Amdahl 5680 computer. The Liverpool group took advantage of improvements in computer technology and GIS and ran their package on a Pentium PC with 32Mb of RAM. In an attempt to address the Modifiable Areal Unit Problem (MAUP), the analysis was repeated for many different sized circles as it was believed that real clusters would be apparent at the same location at every spatial scale. The profiler is a more complex operation and includes a number of socio-economic data sets which may not be available to all researchers or police officers.

8.3. EVALUATION OF CURRENT CRIME-SPECIFIC HOTSPOT PROGRAMS

Two popular methods for displaying crime hotspots involve using a program called STAC, and Vertical Mapper – an add-on program for MapInfo. The program for investigating the Spatial and Temporal Analysis of Crime (STAC) has been designed by the Illinois Criminal Justice Information Authority based in Chicago, and is popular amongst US crime analysts. A third option, recently available, will also be discussed in this section.

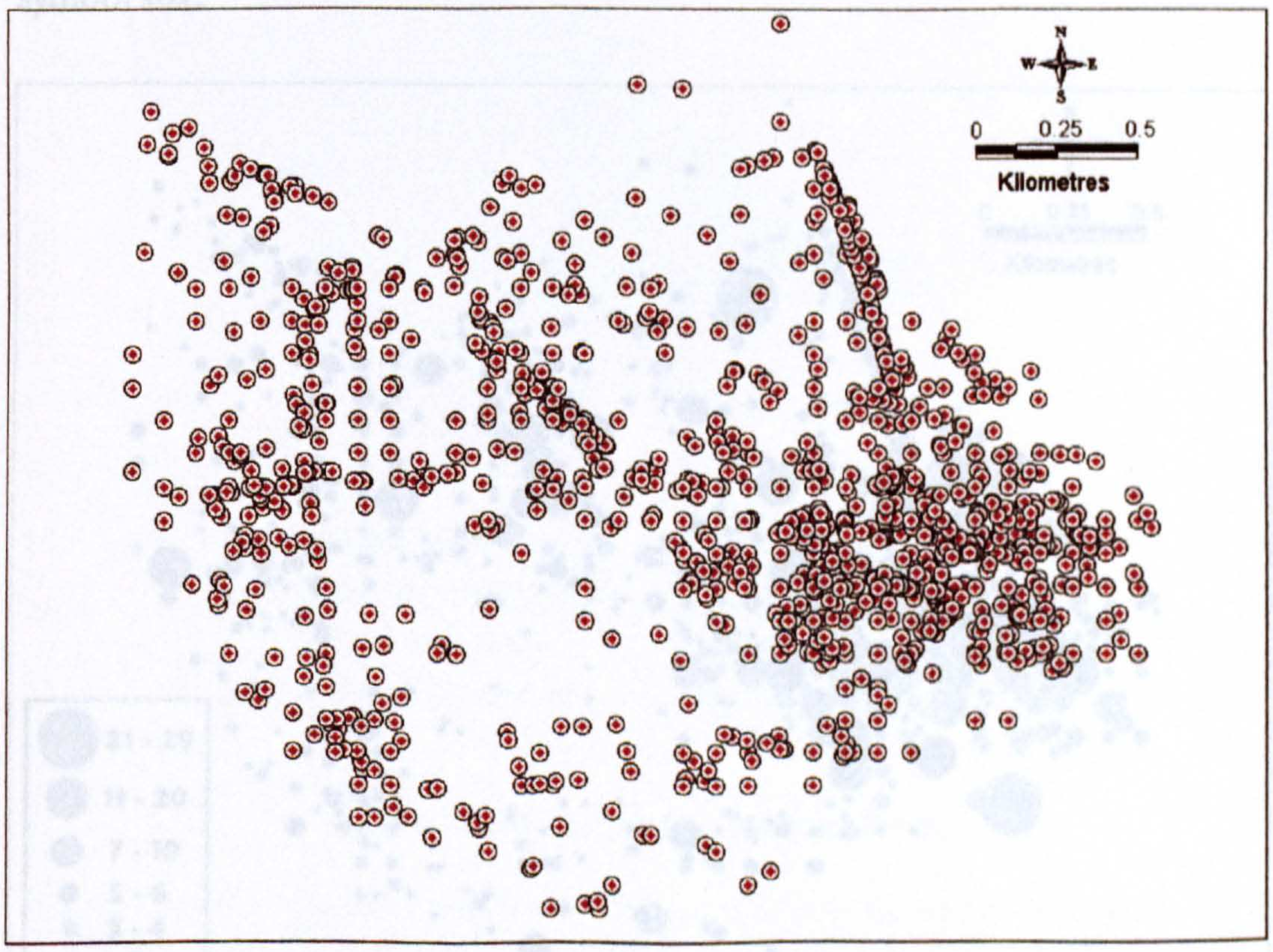


Figure 8-8 Central division non-residential burglaries (April 95 to April 97).

Throughout this section of the chapter an example data set is used drawn from non-residential burglaries in Nottingham Central division over the period April 1995 to April 1997. Central division contains the Central Business District (CBD) for Nottingham and has a high concentration of shops and offices. In the two year period there were 2524 non-residential burglaries. These are shown in Figure 8-8 which also displays one of the limitations of point mapping with a standard GIS. Although the image does show exactly where the crimes occurred, there is a problem of overlapping points. Crime events which occur close to each other will merge into one as the graphical resolution deteriorates. A second problem with this type of display is the visualisation of repeat victimisation. Although the GIS is creating a symbol for each crime incident, whenever there is repeat victimisation the subsequent symbols are placed in exactly the same place as the preceding points. One symbol can indicate either one crime or a number of incidents. This is the case in the non-residential burglary picture of Central division where the highest number of incidents at one location over the two year period was 29. This repeat victimisation effect is shown in Figure 8-9 where the number of reported burglaries occurring at each location is indicated by the symbol size.

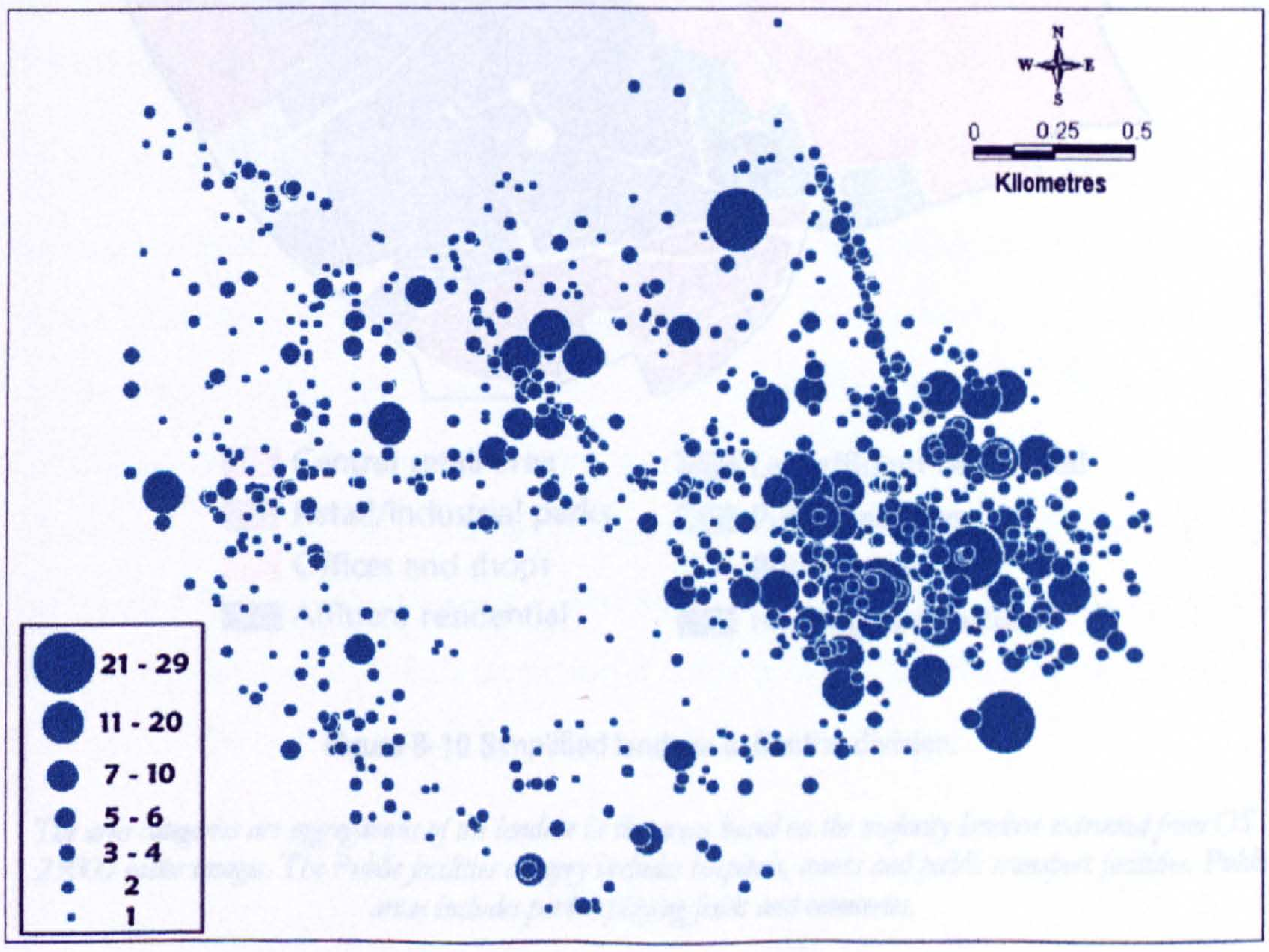


Figure 8-9 Non-residential burglaries. Symbol size indicates number of reported crimes.

For clarity and confidentiality the following images only indicate the crime locations without the underlying city structure. Although this emphasis allows a clearer view of the crime locations and the hotspot structures, the sites of the non-residential crime events are related to the urban structure. Figure 8-10 shows a simplification of the landuse pattern of Central division for reference.

Central division has a central shopping district (purple shading) surrounded by offices and public facilities. To the immediate west of the city centre is an affluent area of expensive housing known as The Park (dark blue shading in Figure 8-10). The rest of the division is made up of poorer housing to the north-west (Radford) and student accommodation in the south-west (Lenton).

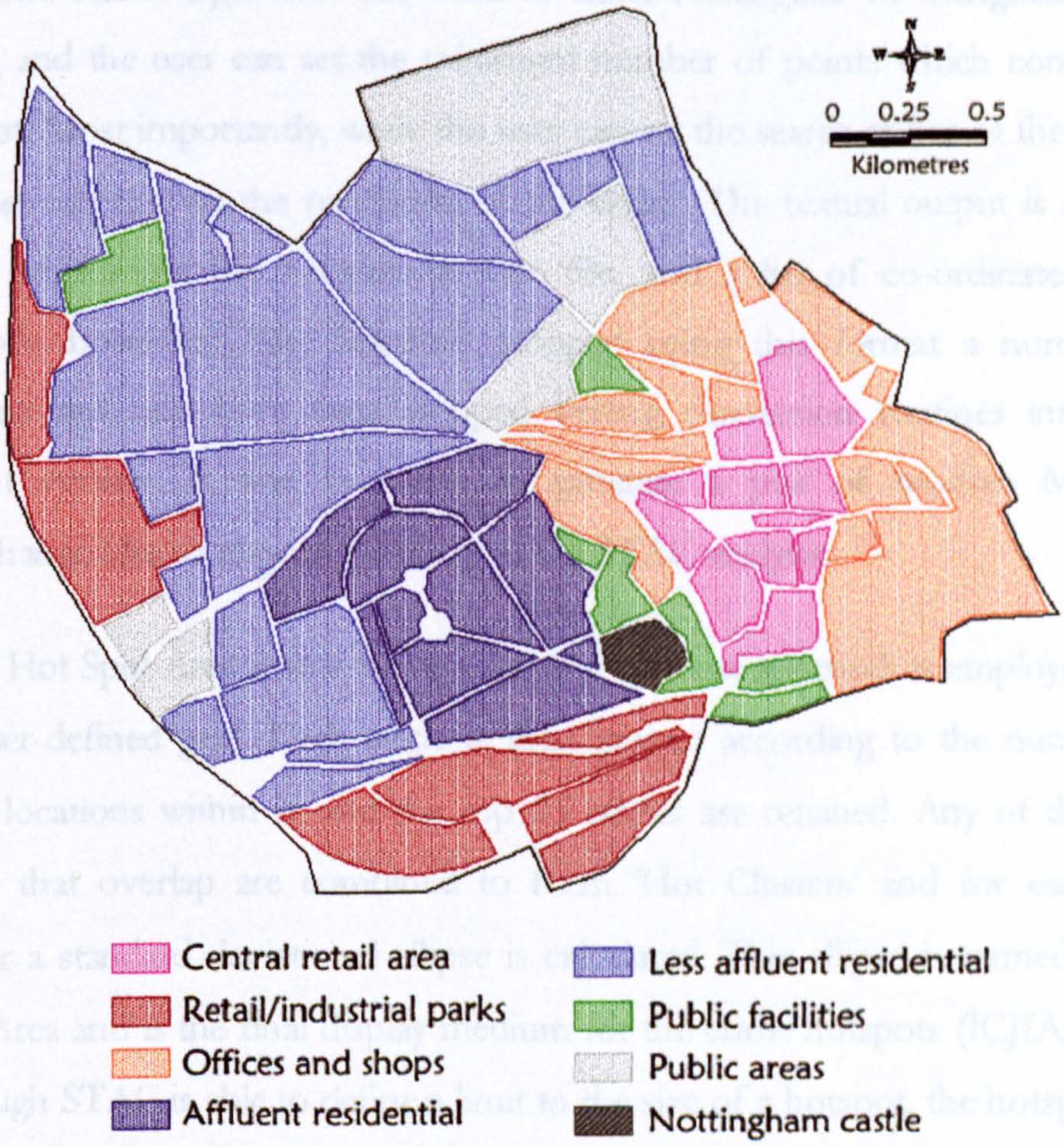


Figure 8-10 Simplified landuse in Central division.

The area categories are aggregations of the landuse in the areas based on the majority landuse extracted from OS 1:25000 raster images. The Public facilities category includes hospitals, courts and public transport facilities. Public areas includes parks, playing fields and cemeteries.

8.3.1. STAC

STAC is freely available from the Illinois Criminal Justice Information Authority and is a suite of DOS based routines designed, amongst other things, to detect clusters in co-ordinate georeferenced crime data. The program employs a moving window approach with a selectable circle bandwidth. It is a little dated now (current issue is version 4.0) and does pose some problems in getting a comprehensible output. The program takes in point data as a comma delimited values file. A number of basic program functions are adjustable through a modifiable parameters file. These parameters include the ability to change the grid upon which the analysis takes place and to window a portion of the study area. The search algorithm can work in either rectangular or triangular search mode, and the user can set the minimum number of points which comprise a hotspot. Most importantly, while the user can set the search radius of the circles, the user can not set the resolution of the lattice. The textual output is a set of mean centres for the hotspots in one file, and a list of co-ordinates for a MapInfo Boundary File. MapInfo stopped using this format a number of versions ago, and have even stopped writing conversion routines into their current version. It was necessary to generate a pair of modern MapInfo Interchange files to import the result of the STAC analysis.

In the Hot Spot Area search mode a moving window approach is employed over the user-defined grid. Each circle is then ranked according to the number of crime locations within it, and the top 25 circles are retained. Any of these 25 circles that overlap are combined to form 'Hot Clusters' and for each Hot Cluster a standard deviational ellipse is calculated. This ellipse is termed a Hot Spot Area and is the final display medium for the crime hotspots (ICJIA, 1996). Although STAC is able to define a limit to the size of a hotspot, the hotspots are always displayed as standard deviational ellipses, a shape which generally bears little resemblance to the underlying crime morphology.

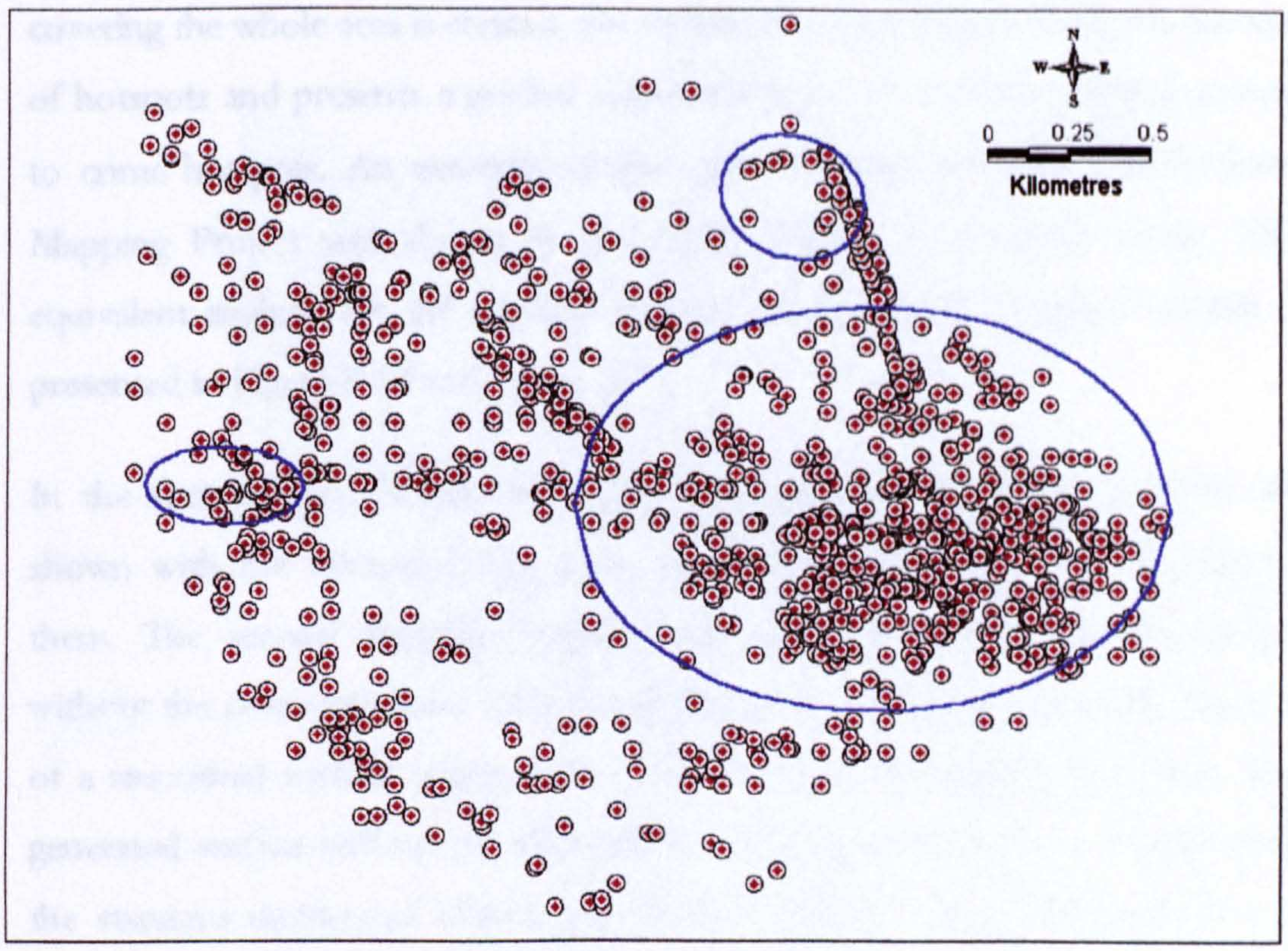


Figure 8-11 Central division non-residential burglaries for April 1995 to April 1997, with STAC hot spot areas.

STAC generated 'hot clusters' are shown as blue ellipses.

Figure 8-11 is an example of the type of output which STAC produces. The co-ordinates of every non-residential burglary on Nottinghamshire Constabulary's Central division from April 1995 to April 1997 (2524 recorded crimes) were used to generate the hotspots outlined in blue. As can be seen from this image, the program has identified three hotspots: two small hotspots to the North and West of the city centre, and a main hotspot which covers the whole of the city centre.

8.3.2. Vertical Mapper

Vertical Mapper™ for MapInfo® is a software package which allows for the analysis and display of continuously varying data (Northwood Geoscience, 1998), and is essentially a raster manipulation package for MapInfo. Vertical Mapper is able to perform basic surface modelling and DEM construction which can create a surface of criminal activity. This feature has been used for crime mapping in the London Borough of Brent (BCMP, 1998) though while a surface covering the whole area is created, this surface does not impose limits on the size of hotspots and presents a gradual increment from areas of no criminal activity to crime hotspots. An example of this type of image from the Brent Crime Mapping Project was shown in an earlier chapter (2: Previous work). The equivalent analysis for the Central division non-residential burglary analysis is presented in Figure 8-12 and Figure 8-13.

In the first diagram (Figure 8-12) the triangulation interpolated hotspots are shown with the corresponding crime locations which were used to generate them. The second diagram (Figure 8-13) shows the same hotspot surface without the crime locations. This clearer image shows that the hotspots are part of a smoothed surface which more closely follows the pattern of crimes. The generated surface follows the flow of the crime patterns more accurately than the standard deviational ellipses produced by STAC. These hotspots are not delimited in any way, and placing a limit on the extent of a hotspot from the total surface would be entirely subjective.

It must also be understood that Vertical Mapper is neither inexpensive nor easy to use. The program does not guide the user through the process and it does not clearly state the best way to achieve the desired output. A reasonable degree of user experience is expected.

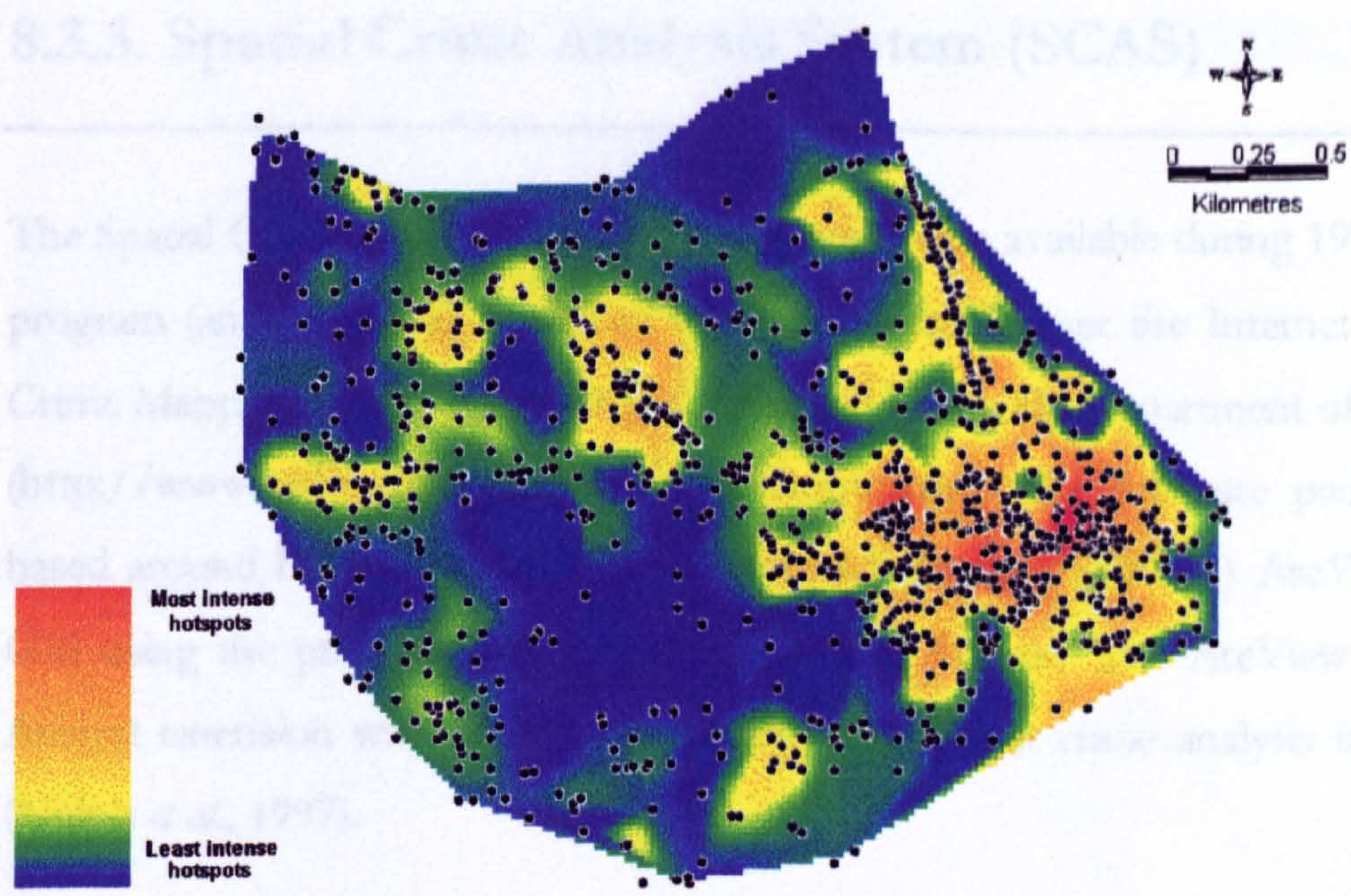


Figure 8-12 Central division non-residential burglaries (shown as black dots) with Vertical Mapper hotspot surface

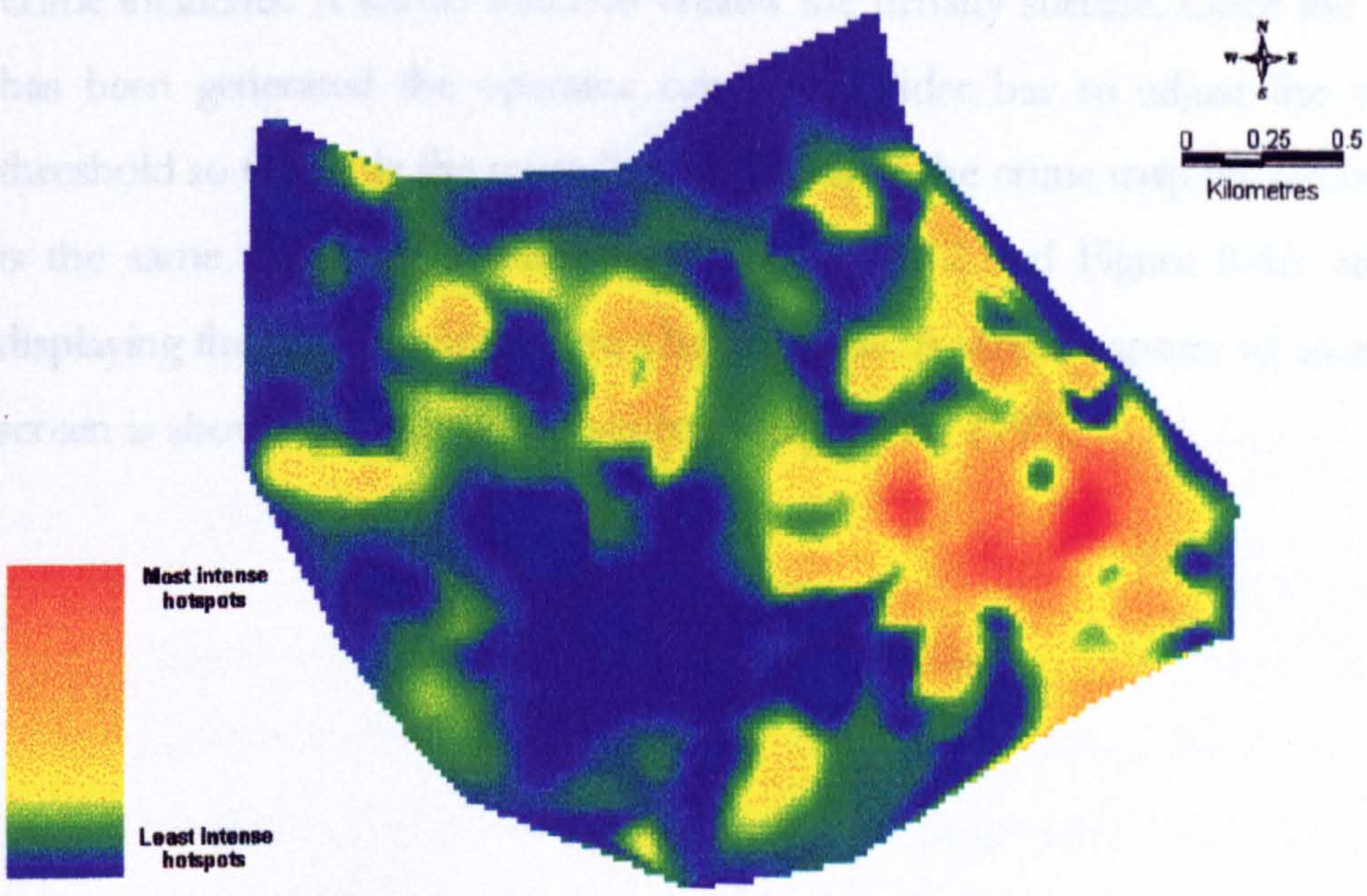


Figure 8-13 Central division non-residential burglary hotspot surface only (generated by Vertical Mapper)

8.3.3. Spatial Crime Analysis System (SCAS)

The Spatial Crime Analysis System (SCAS) was made available during 1998. The program (an add-on for ArcView) was offered freely over the Internet by the Crime Mapping Research Center, a department of the US Department of Justice (http://www.usdoj.gov/criminal/gis/scashome.htm). The software package is based around Environmental Systems Research Institute's (ESRI) ArcView 3.0 GIS using the programming interface language Avenue. The ArcView Spatial Analyst extension was also used to develop the spatial crime analysis modules (Nulph *et al.*, 1997).

The software offers two options for 'hotspot' analysis; surface-derived hotspots and standard deviational ellipses. These are essentially the same options provided by Vertical Mapper and STAC. The surface-derived hotspots are generated using ArcView Spatial Analyst and builds a surface of incident density for the chosen crime incidents. A kernel function creates the density surface. Once the surface has been generated the operator can use a slider bar to adjust the visibility threshold so that only the more 'intense' areas of the crime map are shown. This is the same as adding a threshold to Figure 8-12 and Figure 8-13, and only displaying those areas which have a higher value. A screen capture of an example screen is shown in Figure 8-14 (source: Nulph *et al.*, 1997).

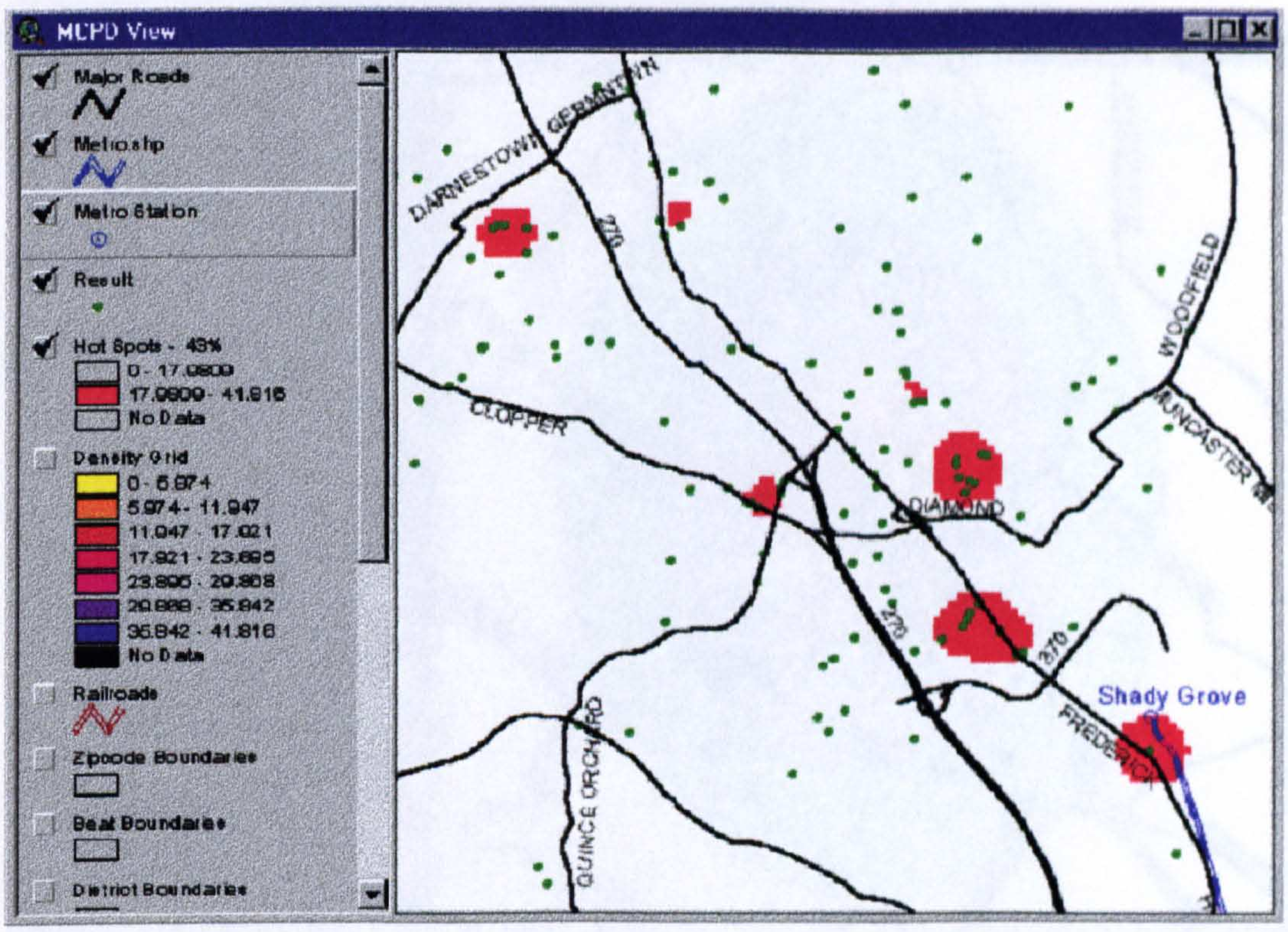


Figure 8-14 Screen capture from SCAS surface generated hotspot routine.

The standard deviation ellipse routine is actually cruder than STAC and simply generates a global standard deviational ellipse based on all the available crime data in the current query. Geometric and statistical properties of standard deviational ellipses can be found in a variety of texts (for example Ebdon, 1996). The value of a single global standard deviational ellipse to spatial crime analysis and hotspot analysis is not made clear by the authors of SCAS, though they do say;

> *Regions based on standard deviation are often used in crime analysis in predictive models for serial crimes; if an analyst has determined that a series of crimes was likely committed by a single person or group, known crime locations can be used to help predict where and for what purpose law enforcement personnel can be deployed (Gottlieb et al., 1994; quoted in: Nulph et al., 1997).*

A screen capture from the technical reference paper for SCAS is shown in Figure 8-15.

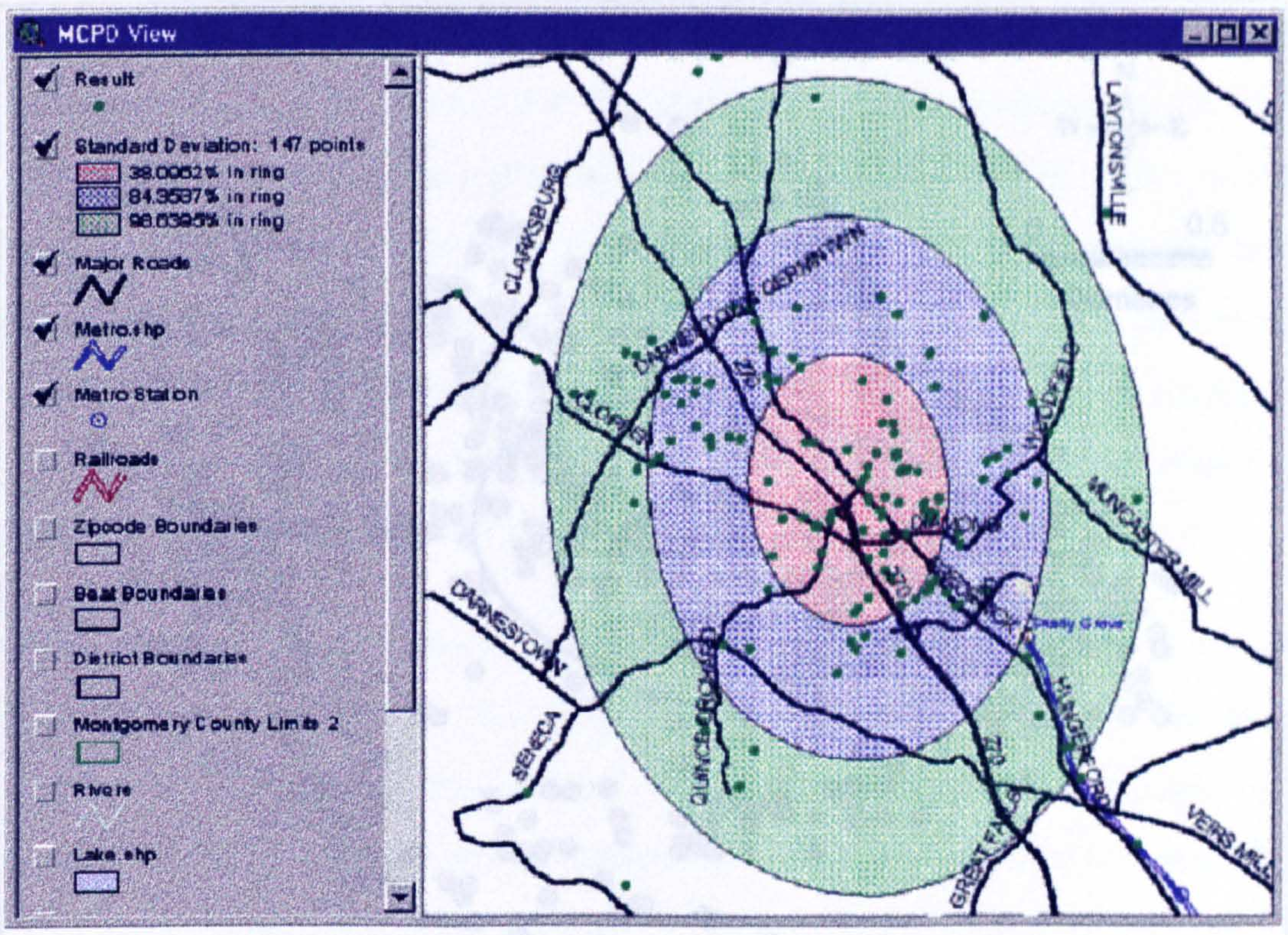


Figure 8-15 Screen capture from SCAS standard deviational ellipse routine.

When this approach is applied to the Central division non-residential burglary data (the example data used throughout this section) the following output is the result (Figure 8-16). As can be seen from this diagram, the standard deviational ellipse is too crude to elicit any useful information and is a retrograde step from using STAC.

SCAS has now been withdrawn from the CMRC web site.

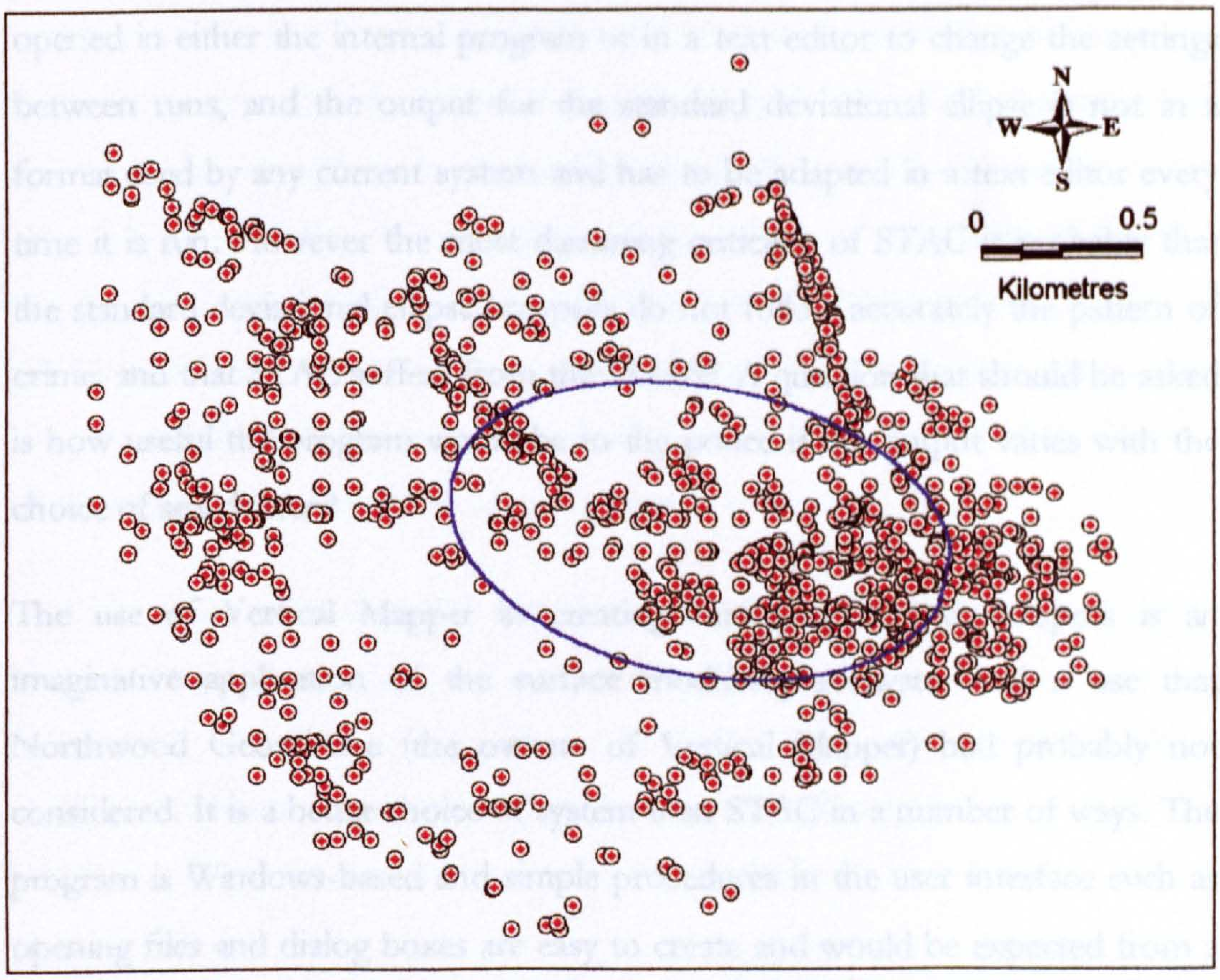


Figure 8-16 Global standard deviational ellipse applied to Central division non-residential burglary data.

8.3.4. Evaluation

Evaluation of the various options is dependent on the time available for the analysis, the capabilities of the user, and the type of output required. In a previous review of STAC and GAM, the GAM was seen as more favourable than the STAC system to the academic user concerned with data quality, statistical significance and methodology. The STAC might be more useful to the police as it provides a simpler graphical output (Hirschfield *et al.*, 1997). For both the police and the academic user, the GAM is currently unavailable, though Openshaw and colleagues are in the process of making the GAM available through the Internet (Openshaw *et al.*, 1998).

STAC is a simple system to use and the output is fairly comprehensible to most users. However it does suffer from a number of limitations. It is a DOS-based program and is less easy to use than a more modern Windows-based system. This is especially the case when opening files. The parameters file has to be opened in either the internal program or in a text editor to change the settings between runs, and the output for the standard deviational ellipse is not in a format used by any current system and has to be adapted in a text editor every time it is run. However the most damming criticism of STAC is probably that the standard deviational ellipse hotspots do not follow accurately the pattern of crime, and that STAC suffers from the MAUP. A question that should be asked is how useful the program would be to the police if the output varies with the choice of search area?

The use of Vertical Mapper in creating surface generated hotspots is an imaginative application of the surface modelling software, and a use that Northwood Geoscience (the owners of Vertical Mapper) had probably not considered. It is a better choice of system than STAC in a number of ways. The program is Windows-based and simple procedures in the user interface such as opening files and dialog boxes are easy to create and would be expected from a modern system. The mechanics of actually developing a crime density surface are however not so easy. There are a number of different interpolation options with which the user is presented with, and each option contains a barrage of adjustable features and possibilities. This means that the user of this type of system should have a good background in digital elevation and terrain modelling to understand the implications of each option in order to choose the correct one. The Brent Crime Mapping Project has been one of the main users of Vertical Mapper for crime mapping (BCMP, 1997; BCMP, 1998), though moving to version 2.0 from version 1.5 caused them some problems. The loss of their GIS manager (Spencer Chainey) to another local authority meant that they have been unable to achieve the same results as before. They feel that this is because Mr. Chainey's background was in GIS whilst theirs is in crime, and the Vertical Mapper interpolation functions require a thorough knowledge of GIS and terrain modelling (PC Mark Patrick – personal communication).

The final point to consider with the use of Vertical Mapper is the quality of the derived surfaces themselves. The derived surfaces cover the whole study area so the user is able to see the distribution of crime density across the whole region. This is certainly advantageous when compared with the STAC system as the user is able to see what is happening outside the main hotspot centres and examine the process across the whole area. The shape of the hotspots also follows better the crime density and more reasonably mimics the underlying surface of the urban crime geometry. What STAC can do, which is unavailable in Vertical Mapper, is to express a limit to the size of a hotspot. While the composition of the STAC hotspots might be flawed by both the shape of the hotspot and the use of a standard deviation as the limiting factor, the imposing of limits does allow for a simple binary classification of hotspots (hot and not).

SCAS is similar to both STAC and Vertical Mapper. Designed for ArcView as Vertical Mapper is designed for MapInfo, it provides much of the same functionality and has the same flaws. The SCAS standard deviational ellipse routine is of limited value, and if standard deviational ellipses are required, the STAC system produces more reasonable ellipses for crime hotspots.

So while each system is able to contribute a useful aspect towards the mapping of crime hotspots (Vertical Mapper's global surface which follows the crime density, STAC's hotspot limits), no one system provides the sought-after solution of a statistically supportable binary classification of hotspots which follow the underlying geometry of the crime distribution.

In an attempt to provide a simple classification of hotspots which are not limited to rigid geometric shapes this study suggests employing both a moving window approach and a Local Indicator of Spatial Association (LISA) statistic for classification purposes. This two stage process is described in the following sections, starting with the program written by the author to generate the moving window surface: SPAM.

8.4. SPATIAL PATTERN ANALYSIS MACHINE (SPAM)

In this study an adaptation of the GAM has been programmed in Visual C++, and for ease of reference has been called the Spatial Point Analysis Machine (SPAM). The system is designed to read in crime data described as co-ordinate values from a text file. From this file, the maximum and minimum values are calculated, along with basic statistical information about the data set (mean position, standard distance and standard deviational ellipse). These calculations are useful for summarising the data set and describing the variation across the whole region, but do not contribute usefully to any search for variation within the data set and the identification of any clusters or 'hotspots'. The search procedure employed by SPAM is a simplified variation on the GAM as the underlying aim is to improve the ability of the police to identify areas of crime concentration, and therefore the underlying population is less important and not included in the calculation. While including the population density may have academic value to some researchers, the police focus their resources on areas of high crime irrespective of the population density variation.

SPAM uses a moving window approach which allows the user to adjust the dimensions of the grid. The user can also adjust the resolution of the grid mesh so that finer or coarser grids are generated. On execution, the program performs the analysis. The great advantage of Visual C++ is that the program has all the user functionality of a fully-functioning Windows program whilst retaining the speed of calculation of a C routine.

As the program does not consider population density, it confines itself to visiting each intersection on the fine mesh grid and overlaying a number of circles. The intensity of crime events is calculated for every circle and then the program moves onto the next grid intersection. The original GAM and the variation employed by the Liverpool team both used a simple count to sum the number of incidents found within the search circle. This system is satisfactory when population density is also a factor, however this was felt to be less useful for policing purposes. The main policing priority is to focus resources into the areas with most crime intensity irrespective of the population density. If a fine mesh is placed over a collection of points, a sum of the points in each circle will often register the same value for a number of circles if the circles are closely overlapping.

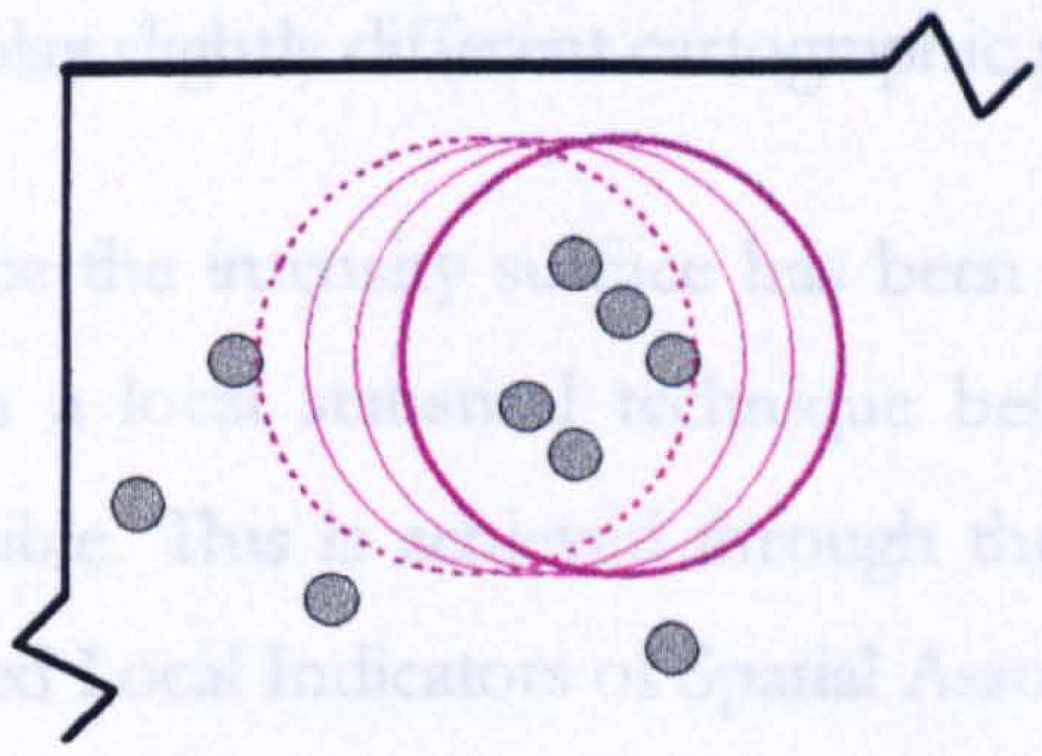


Figure 8-17 An example study area showing four test circles engulfing five spatial events.

Figure 8-17 shows an example where four circles each encompass the same five events. Although the circles are in different places, they will each register the same number of 'hits', even though the events are most centrally located within the thicker test circle. Even though the leftmost circle (dotted) only just contains the points, the total point count is equal to the thick circle. A more informative approach is to measure the intensity of the crime locations relative to their proximity to the centre of the circle by kernel estimation.

SPAM uses a moving window approach with an overlay of 0.8r, intensity values being calculated using the quartic kernel estimation equation (Equation 8-1 on page 223). This process generates the crime intensity surface, which can be seen in Figure 8-18 (Figure 8-19 without the crime points for clarity) for the example data used throughout this chapter (Central division non-residential burglaries). The generated surface is similar to the one produced by Vertical Mapper, except that the cell values indicate a measure of intensity based on an inverse distance weighted estimation.

This surface differs from other surfaces such as the Vertical Mapper output in a number of ways. The process does not aim to interpolate a smooth surface and if there are areas with no crime incidence sandwiched between high crime regions then the no crime area will show a sharp drop in value reflecting the lack of criminal activity. The intensity value for each grid location is a measure of the intensity of the local crime locations. This can mean that a grid site with five nearby locations will register a greater value than a site with seven more remote crime events. Figure 8-18 and Figure 8-19 show the hotspot intensity raster for the Central non-residential burglaries. For clarity the legend in these images has replaced the intensity values for the map classifications with a simpler 'most intense hotspots' and 'least intense hotspots' classification. It should be noted that the MAUP does apply and a different choice of category ranges would display slightly different cartographic products.

Once the intensity surface has been generated it then has to undergo analysis with a local statistical technique before a binary classification of hotspots is feasible. This is achieved through the use of local spatial statistical techniques, called Local Indicators of Spatial Association (LISA).

Two executable versions of SPAM are available on the CD-ROM accompanying this thesis, called **SPAM V1.EXE** and **SPAM V1-1.EXE**. Version 1.1 contains an improved user interface. There is also a help file for the first version (version 1) and two sample data sets (**57.TXT** and **144.TXT**). All of these programs can be found in the **SPAM** folder of the CD-ROM.

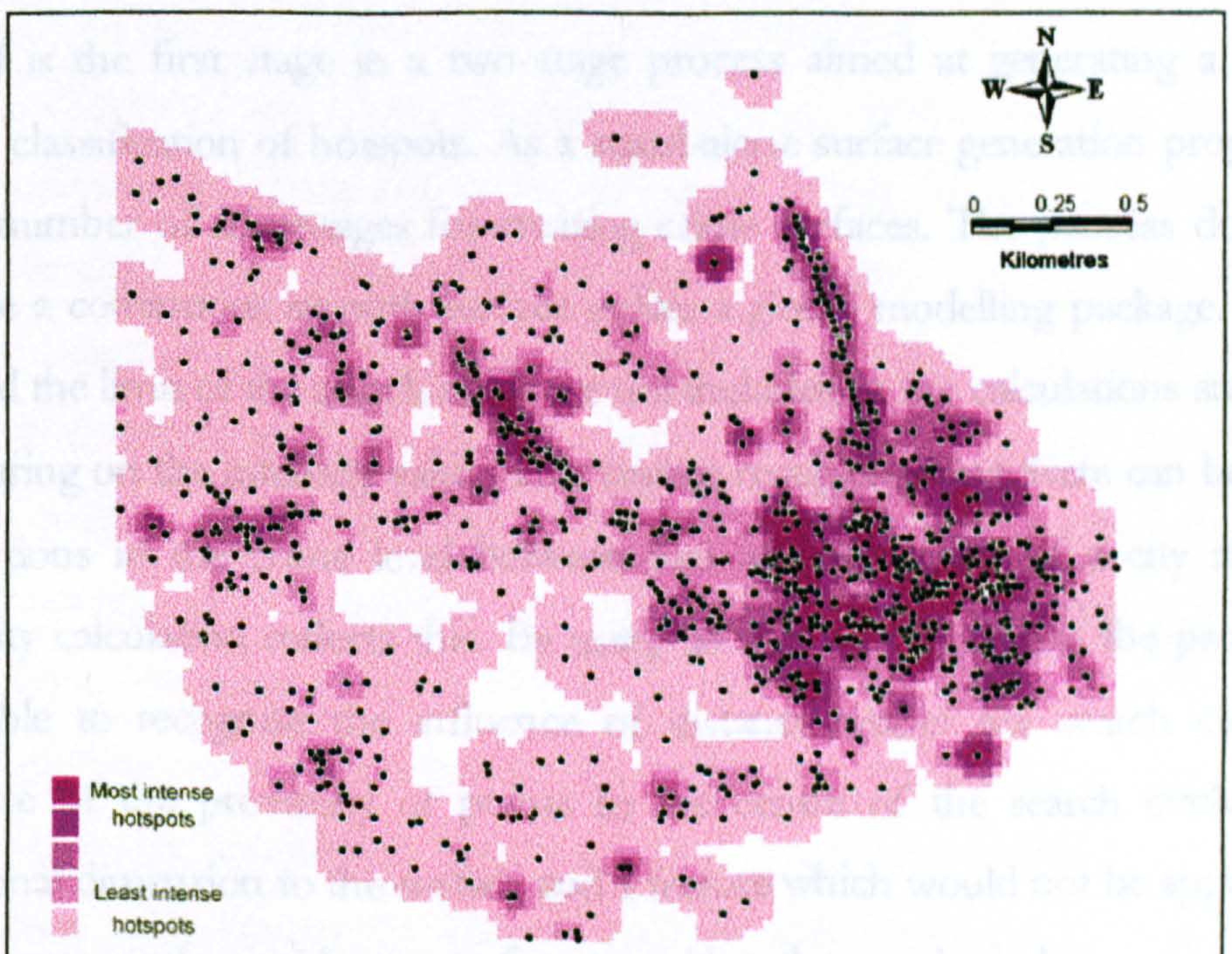


Figure 8-18 SPAM hotspots generated for Central division non-residential burglaries.

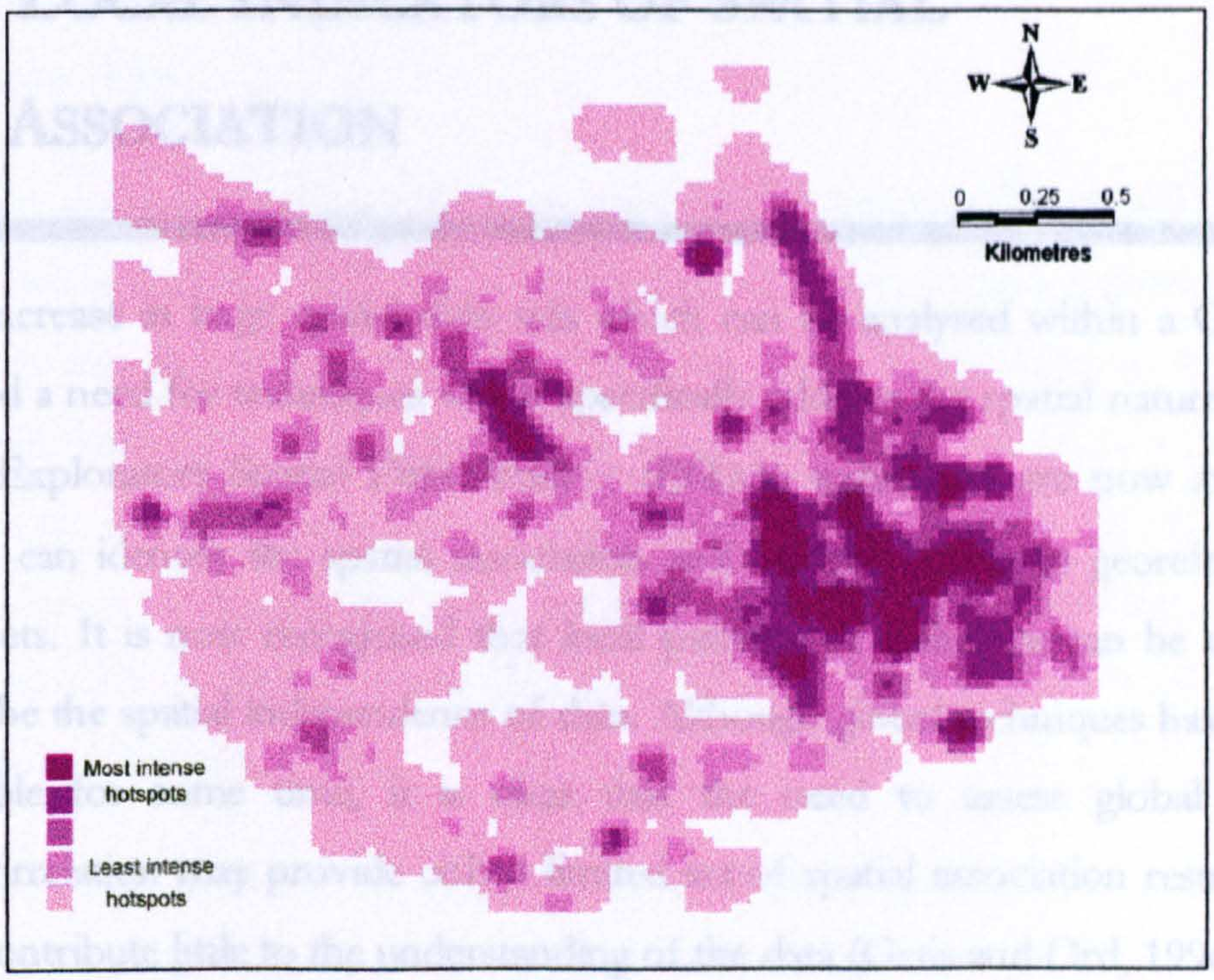


Figure 8-19 SPAM hotspots without crime locations.

EVALUATING SPAM

SPAM is the first stage in a two stage process aimed at generating a simple binary classification of hotspots. As a stand-alone surface generation program it has a number of advantages for creating crime surfaces. The process does not assume a continuous smooth surface unlike a global modelling package. Points beyond the limit of the search circle are not included in the calculations and have no bearing on the intensity value. This system recognises that there can be sharp distinctions in the crime level between two adjacent parts of a city and the intensity calculation reflects this. By using an intensity algorithm the process is also able to recognise the influence of distance within the search circle. A measure of the proximity of points to the centre of the search circle is an additional dimension to the surface and a feature which would not be apparent if the process used a simple count of points within the search circle.

8.5. Local Indicators of Spatial Association

The increase in large spatial data sets which can be analysed within a GIS has created a need for techniques which specifically address the spatial nature of the data. Exploratory Spatial Data Analysis (ESDA) techniques are now available which can identify the spatial association and autocorrelation in georeferenced data sets. It is now recognised that local patterns of influence can be used to describe the spatial independence of data. Although global techniques have been available for some time, it is clear that the need to assess global spatial autocorrelation may provide only a limited set of spatial association results and may contribute little to the understanding of the data (Getis and Ord, 1996).

Local statistical methods have been developed which can assess the spatial association of a variable within a specified distance of a single point. This class of statistic has been growing rapidly in acceptance and many practitioners feel they offer an improvement in understanding the spatial nature of locality (Unwin, 1996). These types of statistic differ from procedures which are applied globally, such as Moran's I and Geary's c which are tested on the complete study area under examination. According to David Unwin (1996), Moran's I and Geary's c can be thought of as a more advanced subset of a larger traditional set of global statistical methods which include nearest neighbour analysis and k-functions. All of these statistical methods can be considered as global in that they require analysis of all or most of the data set and have a null hypothesis that the spatial autocorrelation of a variable is at or near zero. They therefore attempt to characterise the patterning across the entire study region. Unwin (1996) identified three severe limitations in the application of global methods. Firstly, they rely on the assumption that the pattern being tested exhibits stationarity over space. This suggestion that the pattern, and therefore the process that created the variable, is stable is often unreasonable in geography where true areas of spatial homogeneity are rare. An additional factor is that the global methods are vulnerable to edge effects in the calculation, and finally the expectations are often affected by the Modifiable Areal Unit Problem (Bailey and Gatrell, 1995; Openshaw, 1984).

A distinction has been drawn between *general* and *focused* tests (Besag and Newell, 1991; Ord and Getis, 1995), where general tests examine overall patterns across the study area and focused tests concentrate on smaller subset areas of the study area. The latter type of tests have been used when there has been a previous hypothesis about a particular location, such as the location of a nuclear installation as in the example from Besag and Newell (1991). The class of local statistical tests is referred to by the term **LISA** - Local Indicators of Spatial Association (Anselin, 1995).

LISA statistics allow for the decomposition of global statistics and provide an additional tool in exploratory spatial data analysis (ESDA). They can measure dependence in one area of the study region and are particularly adept at: (1) indicating regions of spatial non-singularity, and (2) identifying the existence of local spatial clustering around an individual location, or 'hotspots' (Anselin, 1995; Getis and Ord, 1996).

8.5.1. Definition

Anselin (1995, p.94) suggests that;

> *a* local indicator of spatial association *(LISA) is any statistic that satisfies the following requirements:*
>
> *the LISA for each observation gives an indication of the extent of significant spatial clustering of similar values around that observation;*
>
> *the sum of LISAs for all observations is proportional to a global indicator of spatial association.*

More formally, Getis and Ord (1996, p.262) define a LISA statistic of the form;

$$\Gamma_i = \sum_{ij} w_{ij} y_{ij} \qquad \text{Equation 8-2}$$

where the value of Γ at location i is being determined, w_{ij} represents the spatial association between the target site i and other sites j, and y represents the association of values of a random variable at site i with values at other sites. If both w_{ij} and y_{ij} are elements of matrices X and Y then different methods of constructing the Y matrix will determine the LISA statistic which is employed. Getis and Ord (1996) go on to say that Moran's I statistics can be calculated when y_{ij} is structured as a measure of covariance like the Pearson product-moment statistic, and is of the form $(x_i - \bar{x})(x_j - \bar{x})$, and statistics based on the squared difference - $(x_i - x_j)^2$ - can produce Geary-type statistics such as c_i, K_{1i}, and K_{2i} (Getis and Ord, 1996, p.262).

G_i AND G_i^* STATISTICS

Although all LISA statistics assess the local association amongst the data, the G_i and G_i^* family of statistics evaluate association by measuring additive qualities, and in doing so can compare local averages to global averages. In this manner they are ideal for defining hotspots and placing a spatial limit on those hotspots. The statistics G_i and G_i^*, also written as $G_i(d)$ and $G_i^*(d)$, introduced by Getis and Ord (1992) for the study of local patterns in spatial data, were extended and re-written in 1995 to redefine G_i as a standard variate and to allow for non-binary classifications of the distance d (Ord and Getis, 1995). The initial equation from 1992 is defined below:

$$G_i(d) = \frac{\sum_j w_{ij}(d) x_j}{\sum_j x_j}, j \neq i \qquad \text{Equation 8-3}$$

Although the above equation is more comprehensible and intuitive, the following equations and description are taken from Getis and Ord (1996).

An area is subdivided into n regions, $i=1,2,\ldots,n$, where each region is identified with a point of known Cartesian co-ordinates. Each i has an associated value x_i taken from a variable X. When the focus falls on one particular site i, the values at the remaining observations are denoted as x_j. In a study area of 6 regions (for example) the third site $i=3$ has a value x_3, and the remaining locations (j) have the values x_1, x_2, x_4, x_5, and x_6.

The null hypothesis is that there is no association between the value recorded at site i of variable X, that is x_i and the surrounding values of X, the other values of x_j up to and including a distance d from i. Stated a different way, the null hypothesis is that the sum of values at all the j sites up to a radius d from location i is not more than one would expect by chance given all the values in the study region, including those values beyond distance d.

The null hypothesis appropriate for the $G_i(d)$ statistics require that x_i be excluded from the summation, while the null hypothesis for the statistic $G_i^*(d)$ require that the value at i is included and is summed along with the j values within the distance d if location i. The statistics are formally defined as follows:

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j - W_i\bar{x}(i)}{s(i)\{[((n-1)S_{1i}) - W_i^2]/(n-2)\}^{1/2}}, j \neq i \qquad \text{Equation 8-4}$$

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}^*}{s^*\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}}, all j \qquad \text{Equation 8-5}$$

where $w_{ij}(d)$ is a spatial weights vector with values for all links within d of i. The sum of the weights is defined:

$$W_i = \sum_j w_{ij}(d), j = i \qquad \text{Equation 8-6}$$

and:

$$W_i^* = W_i + w_{ii} \qquad \text{Equation 8-7}$$

$$S_{1i} = \sum_j w_{ij}^2, j \neq i \qquad \text{Equation 8-8}$$

and:

$$S_{1i}^* = \sum_j w_{ij}^2, all j \qquad \text{Equation 8-9}$$

The mean and variance (s^2) calculations are:

$$\bar{x}(i) = \frac{\sum_j x_j}{(n-1)}, j \neq i \qquad \text{Equation 8-10}$$

$$\bar{x}^* = \frac{\sum_j x_j}{n}, \text{all } j \qquad \text{Equation 8-11}$$

$$s(i)^2 = \frac{\sum_j x_j^2}{(n-1)} - [\bar{x}(i)]^2, j \neq i \qquad \text{Equation 8-12}$$

$$s^{*2} = \frac{\sum_j x_j^2}{n} - \bar{x}^{*2}, \text{all } j \qquad \text{Equation 8-13}$$

8.5.2. Characteristics of the statistic

ADVANTAGES

An effect of the use of this statistic to identify the existence of local spatial clustering around an individual location is that it can be used to identify distances beyond which no discernible association exists. Within the analysis of crime hotspots this is a particularly interesting feature with useful benefits. A number of packages are able to produce maps of hotspots but either suffer from the creation of unrealistic hotspot shapes (such as hotspots as standard deviational ellipses in STAC) or create a map by creating a surface of crime concentration without limits to the high crime locations (as in Vertical Mapper). Neither process is particularly justifiable statistically, nor able to set a limit on the size of a hotspot.

LIMITATIONS

In most geographical applications is it unlikely that the underlying distribution of the variable will be normally distributed. Both G_i and G_i^* are however asymptotically normally distributed as *d* increases (Ord and Getis, 1995). Getis and Ord claim therefore that when the number of neighbours is large, approximate normality of the statistics can be assured (1996). They suggest that a conservative choice of d would be such that the number of neighbours (j) is at least 30. They go on to say:

> *when* n *is relatively small, as few as eight neighbours could be used without serious inferential error unless the underlying distribution is very skewed. (p.265)*

8.5.3. Statistical significance

When the test is performed using either the G_i or the G_i^* statistic, there is always the problem of overlap. This is caused when selected sites of i are close together, as is usually the case when the test is performed on every location in the study area (easily done with a computer program). It is also a factor when the choices of distance d are large. It is understandable therefore that sites near each other will display similar values of the spatial statistic, as the search areas of each site i will overlap to some degree. This overlap means that nearby values of i will have correlating values of j, and the greater the overlap, the greater the correlation between the spatial statistic. This autocorrelation between local statistics' values means that they lack independence and it is necessary to use a procedure to sanction the results. Ord and Getis (1995) suggest a Bonferroni test and have produced a table of standard values for a variety of percentiles for the largest G_i or G_i^* of n values. The table is reproduced in Appendix C.

When the G_i^* statistic is applied, the result is a simple binary classification of areas into 'hot or not'. The next two diagrams show the process applied to Central division burglary data, which has been used as the example data throughout this chapter. The 99.9% confidence level was selected for the binary surface after a number of surfaces were constructed from different confidence levels. These different surfaces can be seen in Figure 8-20. The effect of increased statistical rigour is to generate a lower number of small isolated hotspots. The main hotspot areas are not significantly reduced in size, indicating that these areas are still hotspots at a variety of statistical levels. From these different surfaces it was decided that the 99.9% surface gave the most compact hotspot rendering based on visual interpretation and also generated the least number of small isolated point-centred hotspots (hotpoints). A more detailed examination of the structure of hotspot areas can be found in the next chapter.

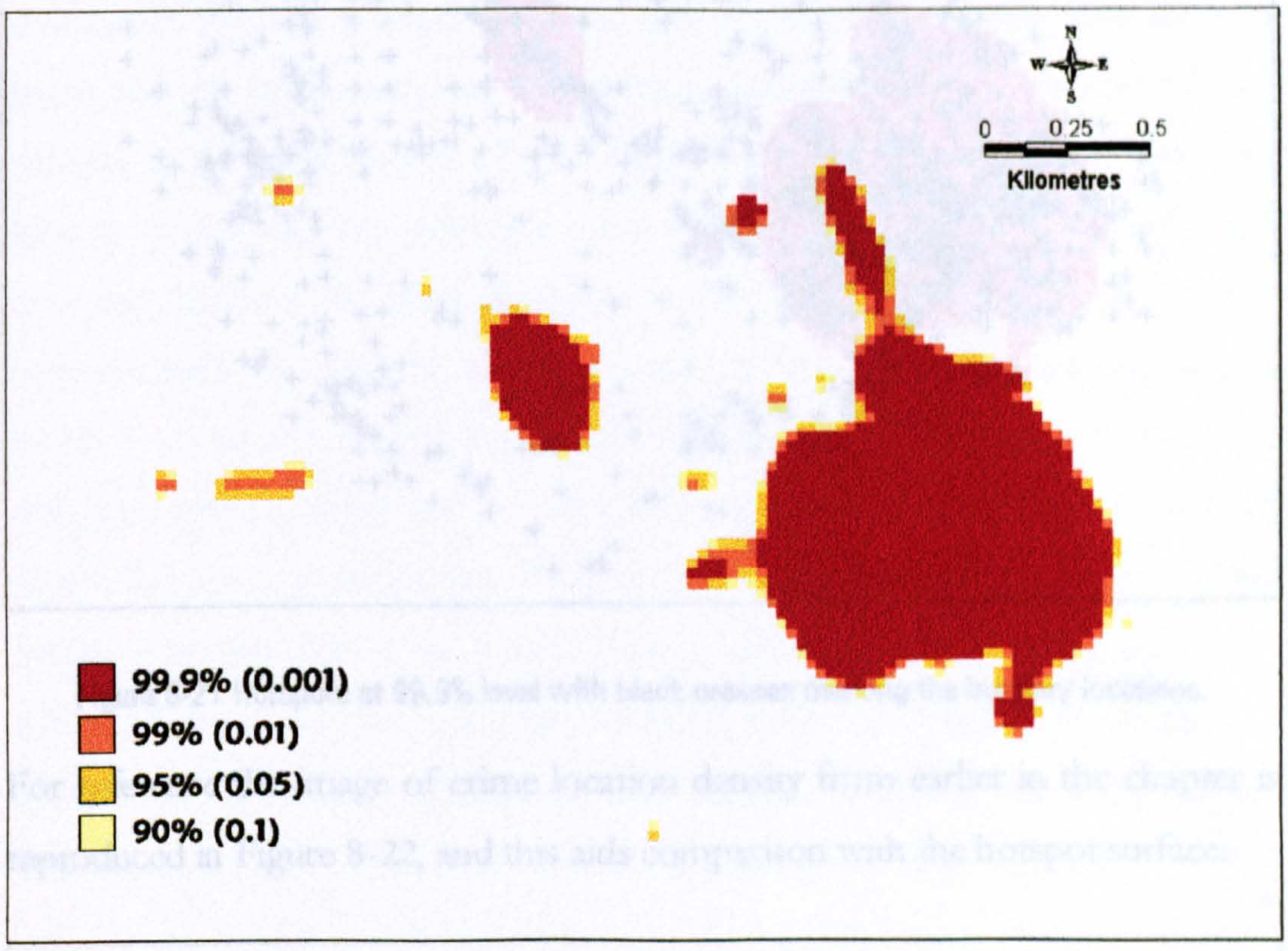


Figure 8-20 Hotspot surface generated by Getis and Ord G_i^* statistic at varying confidence levels.

The second diagram (Figure 8-21) shows the hotspot area when a 99.9% confidence level has been selected and where the crime locations have been changed to simple black crosses to enable the hotspot areas to be view more clearly. More of the underlying hotspots surface can be seen though the crosses do not give any indication of the number of crimes at each location. This type of hotspot has a number of advantages over other hotspot processes. It is simple to understand, follows the architecture of the crime points that generated the hotspots, and is statistically supportable.

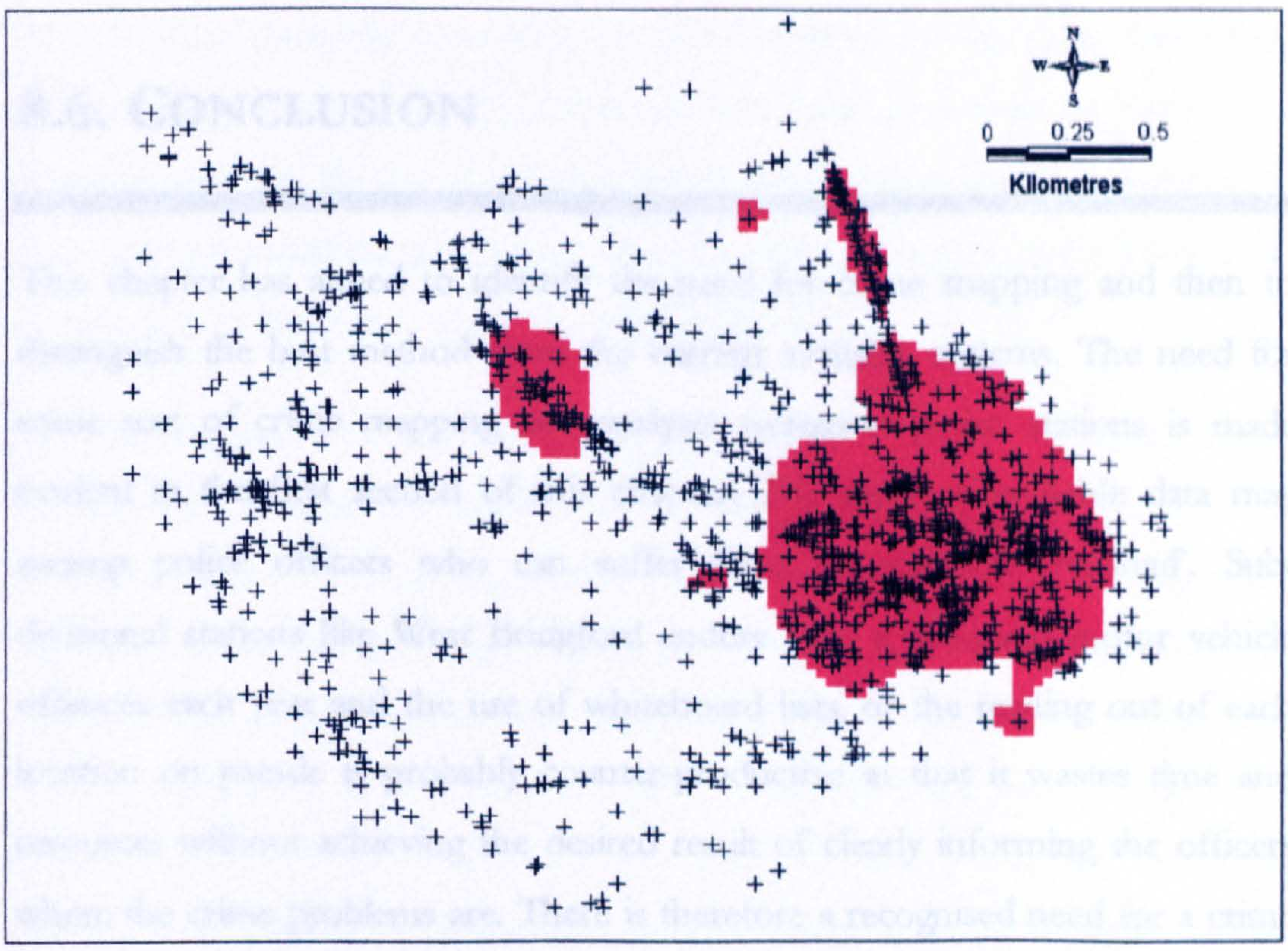


Figure 8-21 Hotspots at 99.9% level with black crosses marking the burglary locations.

For reference the image of crime location density from earlier in the chapter is reproduced in Figure 8-22, and this aids comparison with the hotspot surface.

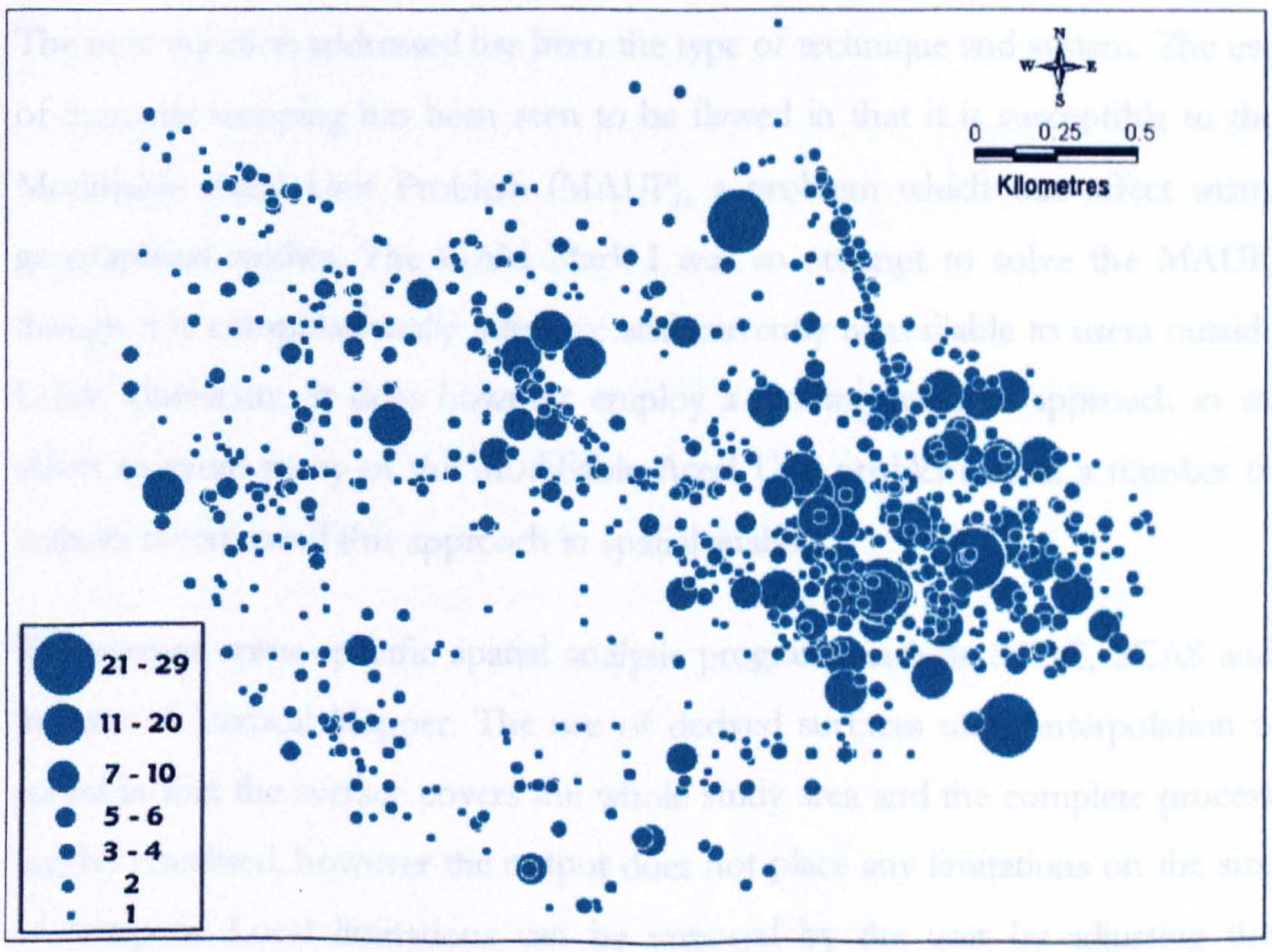


Figure 8-22 Non-residential burglaries. Symbol size indicates number of reported crimes.

8.6. CONCLUSION

This chapter has aimed to identify the need for crime mapping and then to distinguish the best method from the current available systems. The need for some sort of crime mapping and analysis system in police stations is made evident in the first section of this chapter. The mass of available data may swamp police officers who can suffer from 'information overload'. Sub-divisional stations like West Bridgford endure over a thousand motor vehicle offences each year and the use of whiteboard lists, or the reading out of each location on parade is probably counter-productive in that it wastes time and resources without achieving the desired result of clearly informing the officers where the crime problems are. There is therefore a recognised need for a crime mapping system within the police service. Simple point mapping techniques within a GIS are not well equipped to indicate density of points as overlapping symbols are difficult to distinguish and multiple incidents happening at one location are shown on top of each other, appearing as a single symbol.

The next question addressed has been the type of technique and system. The use of thematic mapping has been seen to be flawed in that it is susceptible to the Modifiable Areal Unit Problem (MAUP), a problem which can affect many geographical studies. The GAM Mark I was an attempt to solve the MAUP, though it is computationally intensive and currently unavailable to users outside Leeds University. It does however employ a moving window approach in an effort to avoid many of the Modifiable Areal Unit problems, and a number of authors recommend this approach in spatial analysis.

The current crime-specific spatial analysis programs include STAC, SCAS and the use of Vertical Mapper. The use of derived surfaces using interpolation is useful in that the surface covers the whole study area and the complete process can be visualised, however the output does not place any limitations on the size of hotspots. Local limitations can be imposed by the user by adjusting the display threshold, though no guide is available as to where the threshold should be and this is generally considered to be an exercise for the user. The STAC system uses a moving window approach and does set a limit on the size of hotspots, however the use of one standard deviation seems an arbitrary choice of limit with no statistical significance, and the elliptical output rarely mimics the distribution of the crime data it was derived from.

The solution suggested by this thesis is of a two stage system employing both a moving window approach and the use of a local statistical technique to limit the size of the hotspots. A moving window method using a quartic kernel estimation can produce a surface of crime intensity and then a Getis and Ord G_i statistic can be employed to place statistically justifiable limits on the extent of the most crime ridden areas. It is suggested in this chapter that this two-fold approach achieves the best of all the available options in that the hotspots are statistically valid, show a statistically valid surface, and the hotspots are closely related to the original data and follow the original data's distribution, mimicking the architecture of the crime incidents.

It is proposed here that this approaches a definitive answer to the question of how to identify crime hotspots. The technique uses the moving window to avoid many of the problems associated with the MAUP, is statistically valid and the output is reasonable and appears intuitively correct when viewed with the crime locations. This approach has been employed in a study of police crime hotspot perception and details of this study are contained in the next chapter.

Two important questions arise from this. Even when the process outlined here is applied (as it is in the SPAM program) no guide is available for the selection of 1) a suitable bandwidth, and 2) the maximum search limit for the G_i and G_i^* statistic. These questions are addressed in the next chapter, in the context of a study of the accuracy of police perception of high volume crime.

9. Hotspots and police perception

The previous chapter outlined a technique for describing crime hotspots that follow the morphology of the underlying crime distribution and which are limited in area by a statistical test. This chapter will examine these hotspots in depth by looking at their creation and composition. This is done through the example of a study of police perception of hotspots; a study that raises questions about police information dissemination and crime recording practices.

9.1. INTRODUCTION

The previous chapter identified a methodology for the accurate identification of crime hotspots based on a two-fold approach using a surface generation routine followed by the use of LISA statistics to delineate statistically significant high crime areas. This process goes beyond the possibilities available from the currently available software: STAC, SCAS and Vertical Mapper. The use of these more limited range of software packages has been restricted to the creation of crime hotspots purely for the purpose of mapping crime incidence. The previous lack of a workable methodology for the accurate recognition of crime hotspots has meant that few applications of crime hotspot generation have been proposed beyond simple general mapping. The accurate targeting of crime prevention resources is one of the few alternative utilisations suggested (Ratcliffe and McCullagh, submitted), although there are a number of potential application areas. In addition to the detection of hotspots, the technique promoted in the previous chapter allows the identification of 'coldspots' – those areas devoid of significant criminal activity. The identification of these areas may be useful to local council authorities and academic researchers. Potential applications of hotspot detection include the targeting of crime prevention resources, the improvement of information to the public, more accurate pricing of insurance premiums, and the assessment of police information dissemination strategies. The targeting of crime prevention materials, effort and resources can be considered as an extension to the crime mapping scenario. A GIS user may wish to overlay and compare the crime hotspots with the local authority boundaries or a map of housing occupancy and ownership, however the basic crime hotspots remain used essentially as a direct visual comparison. The pricing of insurance premiums would require the integration of two discrete datasets, both of which are jealously guarded by their owners. Insurance companies and the police service both rightly see their data as valuable property, both in a commercial sense and also in their role as guardian of confidential personal information. While a comparison of crime and incident data would be both interesting and intellectually stimulating, the latter data is unavailable to the author and this type of study falls outside the remit of this thesis. The final application suggested, that of assessing police information dissemination strategies, is possible and is the focus of this chapter.

The integration of information technology (IT) into police services has been discussed elsewhere in this thesis. A force wide introduction of a mapping system is a considerable expense for a police service and similar expenditures elsewhere in most constabularies would normally undergo an evaluation process. So how does one evaluate the success of a mapping system or assess the potential for such a system? It would be a simple procedure to compare the number of crimes before and after the installation of a crime mapping system, except that fluctuations in the rate of recorded crime are rarely the result of changes in policing style alone and are under the sway of a complex mix of factors (for examples of different socio-economic links with crime see Allen, 1996; Bottomley and Coleman, 1976; Cohen and Felson, 1979; Elliott and Ellingworth, 1996).

Crime rate fluctuations occur both before and after the introduction of a mapping system. A reduction in crime rate is not necessarily related to the effect of introducing a mapping system, so it is unreasonable to consider post-introduction patterns to be a direct consequence of the mapping process. A mapping system is designed to disseminate intelligence and a more realistic indicator might be to assess the knowledge of the officers the system is designed to inform. A similar assessment might also be of merit to appraise the necessity and potential configuration of a system prior to procurement.

One justification for introducing a crime mapping system into a police force is often built on the hypothesis that police officers and related crime prevention agencies are not fully conversant with local crime hotspots. As such the introduction of a system is viewed as a mechanism to improve the local awareness of police officers and thereby refine their crime detecting abilities. It is possible however that police officers through their day-to-day activities receive enough information to formulate a realistic impression of the distribution of criminal activity and that a crime mapping system may be an unnecessary expenditure. Police officers form this image of their working area from a variety of sources. A mental picture is created through a combination of formal briefings, informal discussions with colleagues and by assessing the nature and type of work the officer is required to attend, including the reporting of criminal incidents (Klinger, 1997; Reiner, 1997). Criminal activity is usually recorded by police officers when visiting the crime scene, over the telephone for minor offences, or in a minority of cases by detecting the crime themselves (Bottomley and Coleman, 1976). Although an officer will not visit the scene of each incident, informal meetings with colleagues and listening to the radio can keep the officer updated with the most recent and relevant events. Work in the United States has shown how influential these informal meetings, during meal times, on the radio and at incidents, are in affecting a police officer's impression of their beat (Klinger, 1997).

The question therefore arises of whether a police officer's impression of the extent of criminality on his beat is accurate. This chapter examines the hypothesis that police officers can construct an accurate perception of crime distribution from exposure to daily policing practices by comparing crime hotspots generated from the recorded crime data with the perceived local hotspots catalogued from surveys with police officers.

9.2. DERIVING DATA FOR THE PERCEPTION STUDY

The perception study conducted in this chapter required the preparation of two data sets. The generation of accurate crime hotspots using the process described in the previous chapter, and a survey of police officers' estimations of the locations of crime hotspots.

9.2.1. Crime hotspots

Hotspots were generated from Nottinghamshire Constabulary recorded crime data for the period April 1996 to April 1997. The study area covered three force sub-divisions; West Bridgford, the Meadows, and Clifton. Three crime types (residential burglary, non-residential burglary, and motor vehicle crime) were chosen for examination on the basis that they were:

1. force and divisional crime prevention priorities,
2. high volume crimes for the three sub-divisions, and
3. have a high reporting rate.

This last point is worth clarifying. Not all crime is reported and the reasons for the non-reporting of crime have been well documented elsewhere (Biderman and Reiss, 1967; Coleman and Moynihan, 1996; Mayhew *et al.*, 1993). It should be noted that results from the British Crime Surveys show that the selected crime types have amongst the highest reporting rates (Mirrlees-Black *et al.*, 1996).

A moving window type analysis was used in the SPAM program in Chapter 8 (Hotspot analysis). A critique and an example of the type of output from the SPAM program is shown in the previous chapter. A major remaining problem was raised in the previous chapter as to what might be a suitable circle radius (bandwidth).

SELECTION OF A SUITABLE BANDWIDTH

If the grid size is a function of the circle size, the question arises of what size circle to choose. STAC user manuals suggest a search criteria of 220 metres based on Chicago city block sizes (ICJIA, 1996). Suburban Nottingham does not conform to such a regular pattern, and so the question was addressed by asking the officers who completed the hotspot survey. While more details of the survey follow this section, the relevance here is that officers were requested to mark their most important hotspots on each map with a shape that reflected their perception of the extent of the hotspot. Most shapes drawn in the survey were roughly circular, and for each shape a radius was calculated relative to the area of the shape. In this manner, an average radius for each crime type and sub-divisional area based on the officers' perceptions was calculated. These were the radii used in the analysis. Table 9-1 shows the average radii recorded by the police officers in metres and the rounded average used for hotspot analysis.

Table 9-1 Average hotspot radii recorded by police officers, with rounded figures used in analysis shown in brackets.

	Residential burglary	Non-residential burglary	Motor vehicle crime
West Bridgford	346 (350)	175 (175)	392 (400)
Clifton	207 (200)	101 (100)	99 (100)
Meadows	125 (125)	141 (150)	150 (150)

Table 9-1 shows a considerable variation in circle radius, especially between West Bridgford and the two smaller sub-divisions of Meadows and Clifton. Clifton and the Meadows are similar sized sub-divisional station areas, and this is reflected in the similarly sized radii, whereas West Bridgford covers a considerably larger area. It is possible that the differences in radii are the result of the officers being given maps at different scales, or that the different road layout of suburban housing in West Bridgford is reflected in the noticeably larger radii.

MODIFYING THE GETIS AND ORD G_i^* STATISTIC

The analysis proceeded to apply the second part of the process; using a G_i^* statistic to delimit significant areas. Initial use of the LISA statistics described in the previous chapter found that any large regions of the study area with no or little crime tended to force the value of G_i^* up so that every area within range of a crime was considered significant in relation to the test statistic. This is because the calculation made no use of the original boundary file and instead used the Minimum Bounding Rectangle (MBR) of the target points. It was therefore felt that a modification of the G_i^* equation was necessary to combat this problem.

In this study, the denominator has been modified to include only values greater than zero. The nature of urban crime is that much of a study area can escape the effects of crime due to simple geography. If a database of residential burglary is constructed, for example, parklands and the central business district of a city would exhibit a negligible or zero crime level. It is extremely difficult to isolate residential only areas in these types of studies, and it is more practical to widen the study to include the whole region, including the zero crime areas. This does, however, have the effect of reducing the value of the denominator and increasing the test statistic for every region. The inclusion of only those sites that have a positive value isolates the locations where crime occurs. This seems intuitively more sensible as hotspots should stand out against a background of other crime areas, and not against an artificial low value background. This modification (shown in **red**) is reflected in the full equation (given in Getis and Ord, 1996);

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}^*}{s^*\{[(nS_{1i}^*) - W_i^{*2}]/(n-1)\}^{1/2}},\ all\, j,\ x_j \neq 0$$

Equation 9-1

where $w_{ij}(d)$ is a spatial weights vector with values for all cells j within distance d of target cell i, W_i^* is the sum of the weights, S_{1i}^* is the sum of squared weights, and s^* is the standard deviation of the data in the cells. This modification to the original formula has the effect of comparing the hotspot areas with a background of only crime-affected locations. In this manner the test statistic is not affected by the size of the MBR, and while it identifies fewer locations, those locations are the most concentrated hotspots of crime.

At this point the other question raised by the previous chapter should now be addressed: what is a suitable search limit for the Getis and Ord G_i^* statistic?

DECIDING ON A SEARCH LIMIT FOR THE G_i^* *STATISTIC*

Figure 9-1 shows the effect of increasing bandwidth for the G_i^* statistic on residential burglary and motor vehicle crime in the Meadows (April 1996 - April 1997). Although the area of significant hotspot activity increases as the search radius is enlarged, the ratio of crimes to search circle area decreases and this has the effect of diluting the intensity of the hotspots.

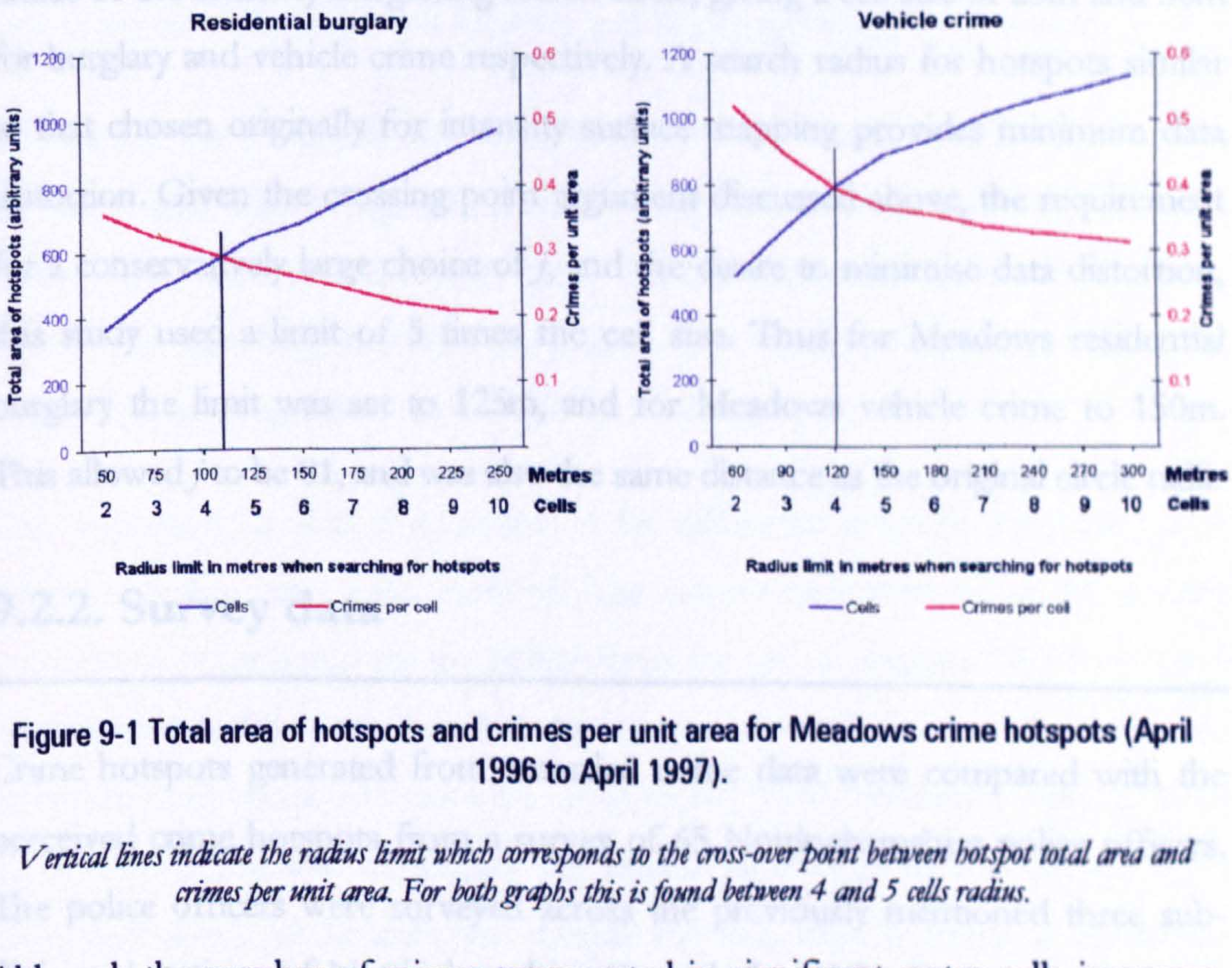


Figure 9-1 Total area of hotspots and crimes per unit area for Meadows crime hotspots (April 1996 to April 1997).

Vertical lines indicate the radius limit which corresponds to the cross-over point between hotspot total area and crimes per unit area. For both graphs this is found between 4 and 5 cells radius.

Although the number of crimes represented in significant raster cells increases, the ratio of crimes to cells decreases, diluting the intensity of the hotspots.

Figure 9-1 shows there is no discernible natural break and the choice of a suitable threshold should be based on a combination factors. Firstly, the crossing point of the two curves could be used as a good indicator of the position of maximum information transfer. Both residential and vehicle crime show this crossing point lies in search radii of between 4 and 5 cells. Converting this to metres means that the search radii should lie in the range 100m to 150m to cover both types of crime. Secondly, the search limit should include sufficient cells to avoid problems of the skewed distributions identified in Getis and Ord (1996). They claim that when the number of neighbours is large, approximate normality of the statistics can be assured, and suggest that a conservative choice of d would be such that the number of neighbours, j, is at least 30. However, they also say that 'when n is small, as few as eight neighbours could be used without serious inferential error unless the underlying distribution is very skewed.' (p.265). A search limit of 3 times the basic cell size would produce $j = 29$ for a G_i^* statistic. Thirdly, consideration must be given to the size of the circle used during the formation of the original intensity surface. The cell size was set to 0.2 of the radius of the intensity integrating search circle, giving a cell size of 25m and 30m for burglary and vehicle crime respectively. A search radius for hotspots similar to that chosen originally for intensity surface mapping provides minimum data distortion. Given the crossing point argument discussed above, the requirement for a conservatively large choice of j, and the desire to minimise data distortion, this study used a limit of 5 times the cell size. Thus for Meadows residential burglary the limit was set to 125m, and for Meadows vehicle crime to 150m. This allowed j to be 81, and was also the same distance as the original circle radii.

9.2.2. Survey data

Crime hotspots generated from recorded crime data were compared with the perceived crime hotspots from a survey of 65 Nottinghamshire police officers. The police officers were surveyed across the previously mentioned three sub-divisional stations of Nottinghamshire Constabulary, UK. Each individual was presented with a number of maps of their patrol area and asked to indicate on these maps the locations of the crime hotspots for the three types of high volume crime: motor vehicle crime, residential burglary and non-residential burglary. The instruction sheets and an example set of blank maps for one sub-division can be seen in Appendix E. Multiple entries were permitted, as were nil returns if the officer felt that there were no areas of concentrated activity. The officers were also asked to ring the hotspot they felt was the most important, and this shape was used as the basis for the hotspot sizes as stated earlier.

Police officers are probably one of the few groups who could be relied upon with some confidence to complete this type of survey. The geographical nature of policing, from county forces down to individual beats, means that an officer has to become intimately familiar with his/her patrol area. On arrival at a new station, officers will more than likely be issued with a local street map to enable them to navigate and perform their duties effectively, and to build up a mental map of the area based on the street map. Reference to the street map becomes unnecessary after a short operational period. In this survey it was found that officers had no problems identifying locations from street maps, even when some officers had not needed to look at a street map for years. The officers were able to visualise mentally each location and the surrounding environment.

The survey was piloted at the Meadows station in February 1998 and extended to Clifton and West Bridgford officers in early March 1998. Readers from large metropolitan areas should bear in mind that these are small police sub-divisions that do not have huge numbers of officers. The Meadows station, for example, has 29 officers assigned to it, of which 26 were surveyed.

One important *caveat* of the research is the difference between the survey time period and the crime data analysed. The respondents were asked to identify hotspots of crime based on the period from the end of August 1997 to the end of February 1998. Data for this period was unavailable so this study has therefore used crime data from the most recent period available, April 1996 to April 1997. A year's worth of data was used in an attempt to iron out any short term fluctuations in the data and identify the long term hotspots as consistent problem areas, areas which it was hoped that the officers would identify in the survey. While the discrepancy in time frames should be noted, it is felt that the more defined hotspots are indicative of long term problems that persisted into the survey time period. This has been confirmed by discussion with officers from all three sub-divisional stations.

By surveying the officers in this manner it was hoped to show the areas where an officer felt there was a centre of intensive criminal activity for different crime types – a hotspot.

9.3. RESULTS

The estimated centroid of hotspot locations from the survey of police officers were digitised and compared with the computer generated hotspots using the 5 cell bandwidths and a G_i^* statistic which accepted locations with a significance level of p=0.001. With a chosen fixed bandwidth it was possible to produce a simple binary classification and categorise the study areas into "hot and not" regions.

The maps on the following three pages show the results of this work. Figure 9-2 shows the Meadows hotspots and survey estimates for residential burglary, non-residential burglary and vehicle crime. Figure 9-3 shows the same maps for Clifton, and Figure 9-4 shows the three crime and hotspot perception maps for West Bridgford. Note the change in scale between the three sub-divisions.

The Clifton hotspot locations appear to be smaller and more numerous, while the hotspots for the Meadows and West Bridgford are larger and less frequent. The shape and characteristics of crime hotspots will be discussed later in this chapter.

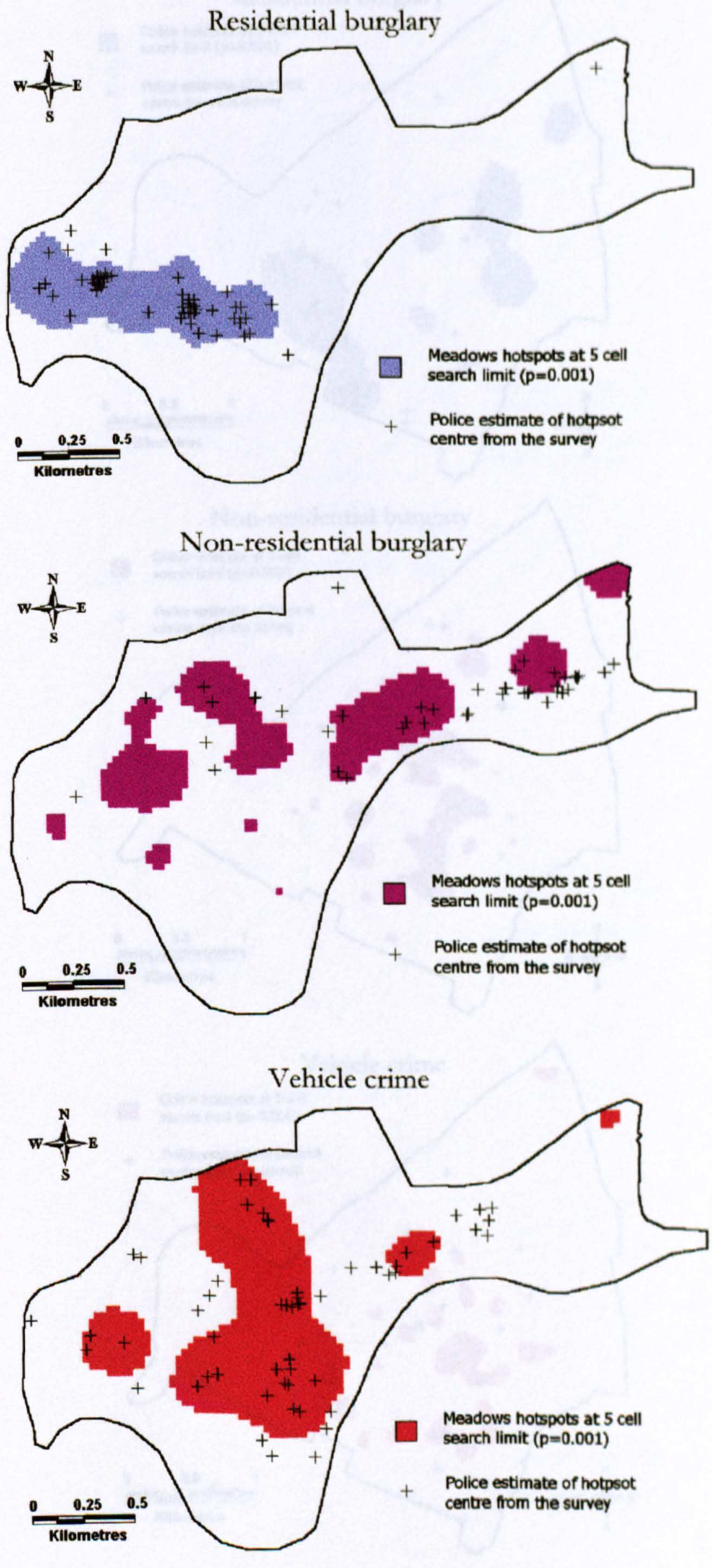


Figure 9-2 Meadows crime hotspots and police estimates.

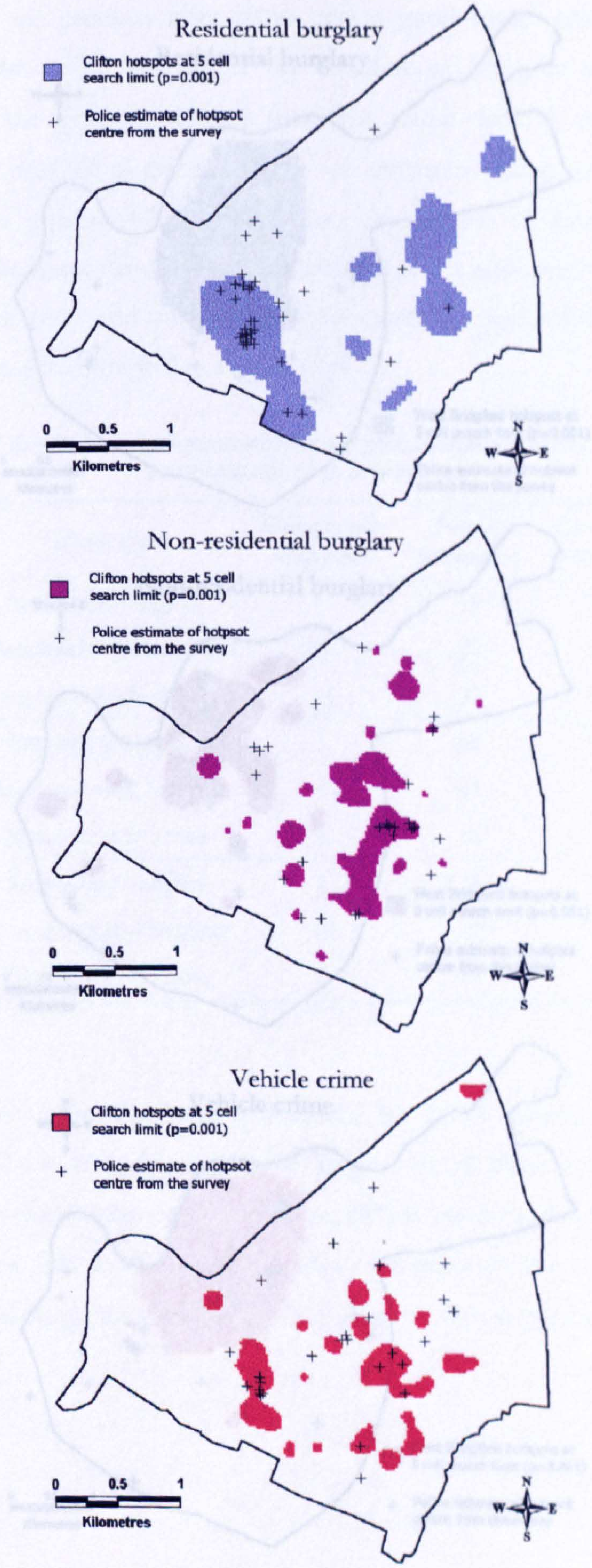


Figure 9-3 Clifton crime hotspots and police estimates.

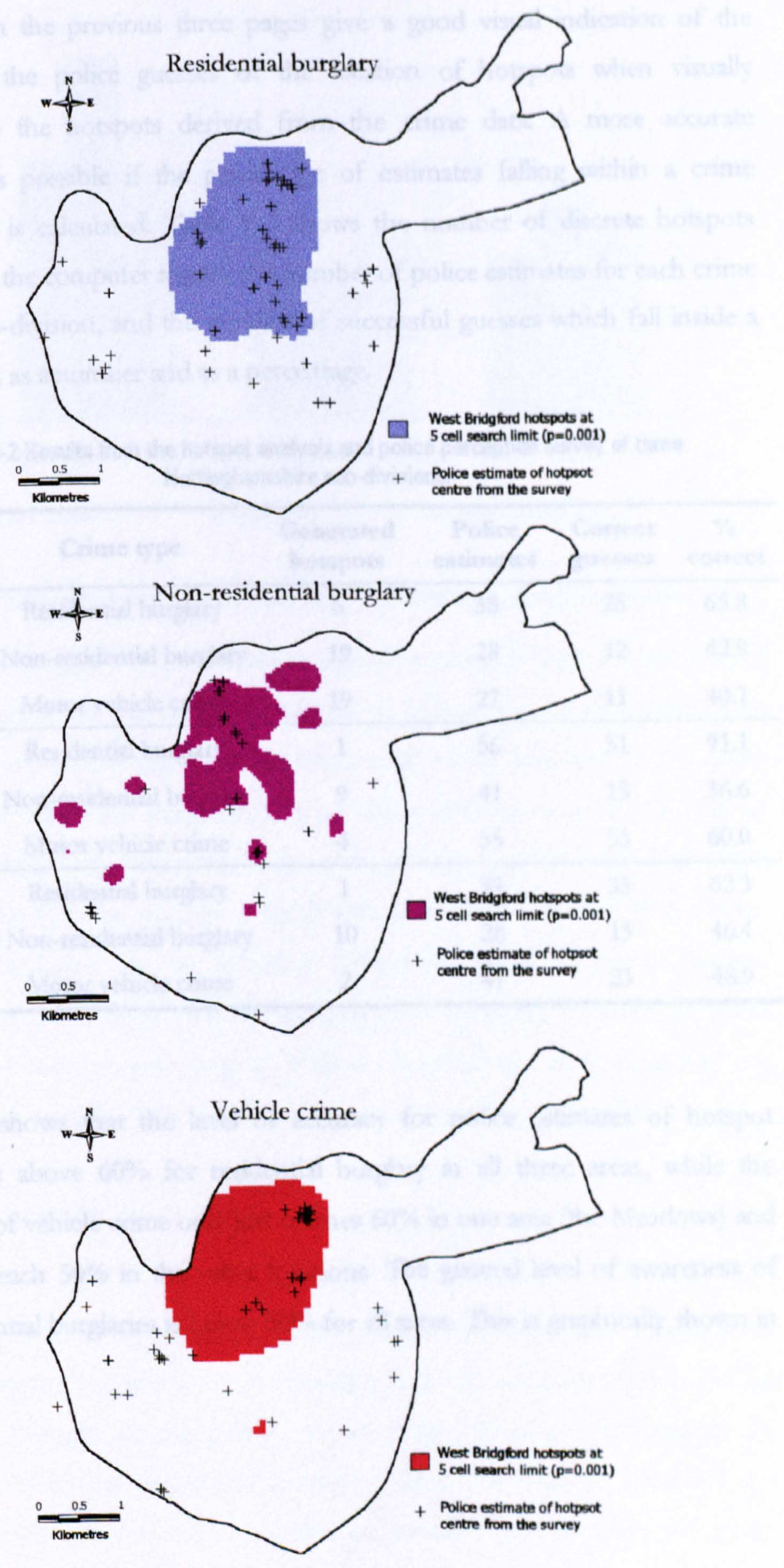


Figure 9-4 West Bridgford crime hotspots and police estimates.

The maps on the previous three pages give a good visual indication of the accuracy of the police guesses of the location of hotspots when visually compared to the hotspots derived from the crime data. A more accurate assessment is possible if the percentage of estimates falling within a crime hotspot area is calculated. Table 9-2 shows the number of discrete hotspots generated by the computer model, the number of police estimates for each crime type and sub-division, and the number of successful guesses which fall inside a hotspot both as a number and as a percentage.

Table 9-2 Results from the hotspot analysis and police perception survey of three Nottinghamshire sub-divisions.

Sub-division	Crime type	Generated hotspots	Police estimates	Correct guesses	% correct
Clifton	Residential burglary	6	38	25	65.8
	Non-residential burglary	19	28	12	42.9
	Motor vehicle crime	19	27	11	40.7
Meadows	Residential burglary	1	56	51	91.1
	Non-residential burglary	9	41	15	36.6
	Motor vehicle crime	4	55	33	60.0
West Bridgford	Residential burglary	1	53	33	62.3
	Non-residential burglary	10	28	13	46.4
	Motor vehicle crime	2	47	23	48.9

Table 9-2 shows that the level of accuracy for police estimates of hotspot locations is above 60% for residential burglary in all three areas, while the awareness of vehicle crime only just reaches 60% in one area (the Meadows) and does not reach 50% in the other locations. The general level of awareness of non-residential burglaries is below 50% for all areas. This is graphically shown in Figure 9-5.

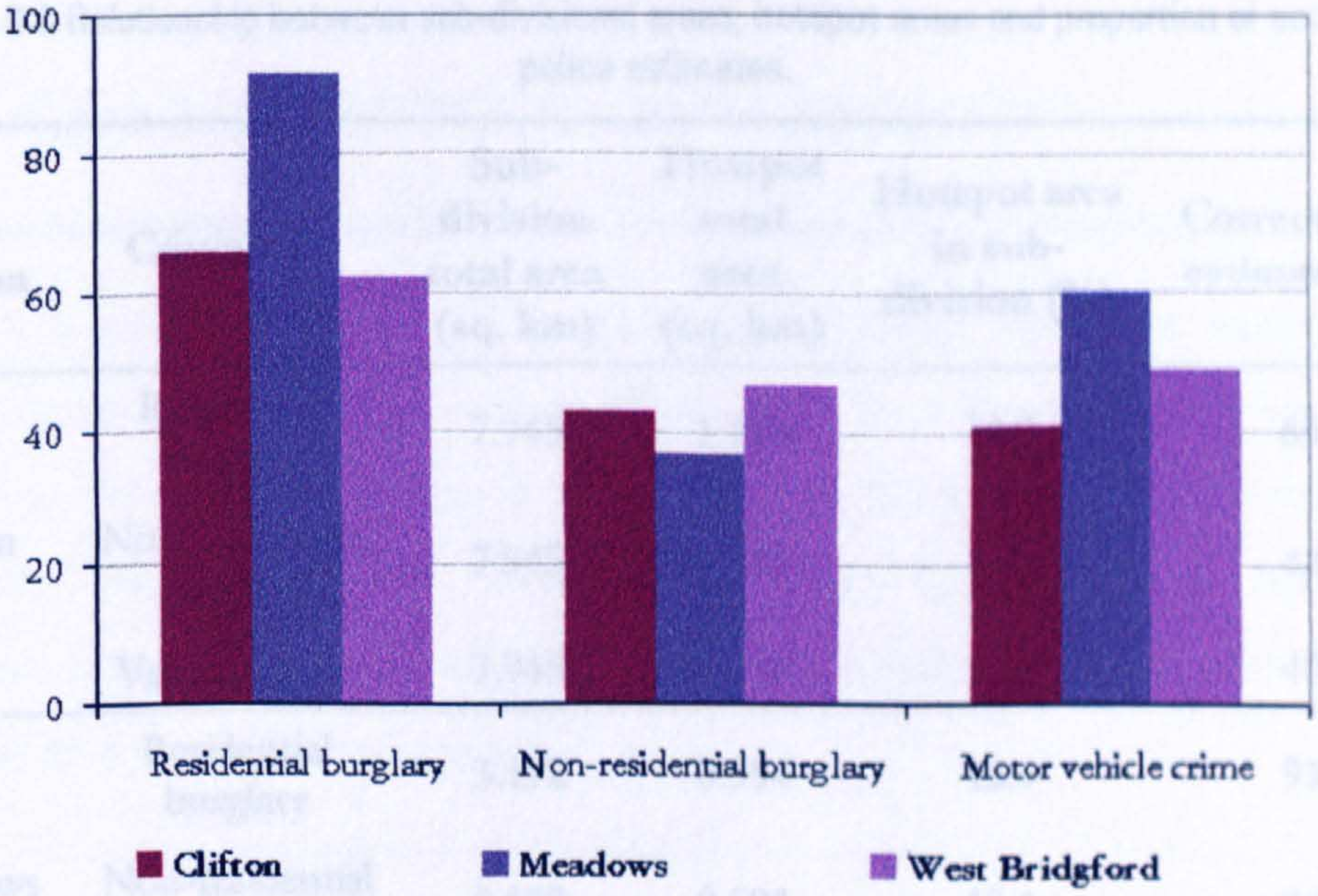


Figure 9-5 Graph showing the success rate of police hotspot perception guesses.

Table 9-2 and Figure 9-5 show that the officers working in the Southern area of Trent division at the sub-divisional stations of Clifton, the Meadows and West Bridgford have a fairly high working knowledge of the locations of the crime hotspots for residential burglary. The Meadows officers in particular had an exceptional knowledge of residential burglary patterns on their 'beat'. This might be due to the fact that the area is small with a smaller population than Clifton and West Bridgford. In the Meadows there are a small number of 'known' burglars and the officers may be locating their estimates around the home addresses of these known burglars. Officers from all the study areas had less success identifying the hotspots of vehicle and non-residential burglary crime.

9.3.1. Hotspot area sizes

The three sub-divisions examined differ in area considerably. The total areas of the hotspots also varies and this is a possible factor in the interpretation (or lack of) of the hotspots by the police officers. Table 9-3 shows the area of the different sub-divisions and the percentage of each area which is identified as a hotspot by the program. The number of accurate guesses from the police perception survey is also shown.

Table 9-3 Relationship between sub-divisional areas, hotspot areas and proportion of accurate police estimates.

Sub-division	Crime type	Sub-division total area (sq. km)	Hotspot total area (sq. km)	Hotspot area in sub-division (%)	Correct police estimates (%)
Clifton	Residential burglary	7.945	1.124	14.2	65.8
	Non-residential burglary	7.945	0.543	6.8	42.9
	Vehicle crime	7.945	0.528	6.6	40.7
Meadows	Residential burglary	3.192	0.394	12.4	91.1
	Non-residential burglary	3.192	0.593	18.6	36.6
	Vehicle crime	3.192	0.797	25.0	60
West Bridgford	Residential burglary	18.73	3.414	18.2	62.3
	Non-residential burglary	18.73	1.957	10.5	46.4
	Vehicle crime	18.73	3.209	17.1	48.9

A number of interesting features arise from Table 9-3. There is a considerable variation in the proportion of each sub-division which is covered by hotspots, ranging from 6.6% (Clifton non-residential burglary) to 25% (Meadows vehicle crime). Although the Meadows residential burglary hotspot covers the smallest area (12.4%) of all the residential hotspots (14.2% and 18.2% for Clifton and West Bridgford), the Meadows officers had the highest success rate (91.1%). When the same Meadows officers were asked about non-residential burglary and had the largest target area to find (18.6% compared to 6.8% and 10.5%) their success rate was the lowest (26.6% compared to 42.9% for Clifton and 46.4% for West Bridgford). This would suggest that at the Meadows station the awareness of local officers for burglary only extends to residential burglaries and not to a knowledge of non-residential crime.

9.4. CHARACTERISTICS OF HOTSPOTS

The following discussion of hotspot characteristics uses the example of residential burglary hotspots in the Clifton area. As in the previous chapter it is useful to view hotspot images in the context of the landuse of the area and a simplified landuse map for Clifton is shown in Figure 9-6. The residential burglary hotspot areas are concentrated within certain residential areas of the sub-division as would be expected.

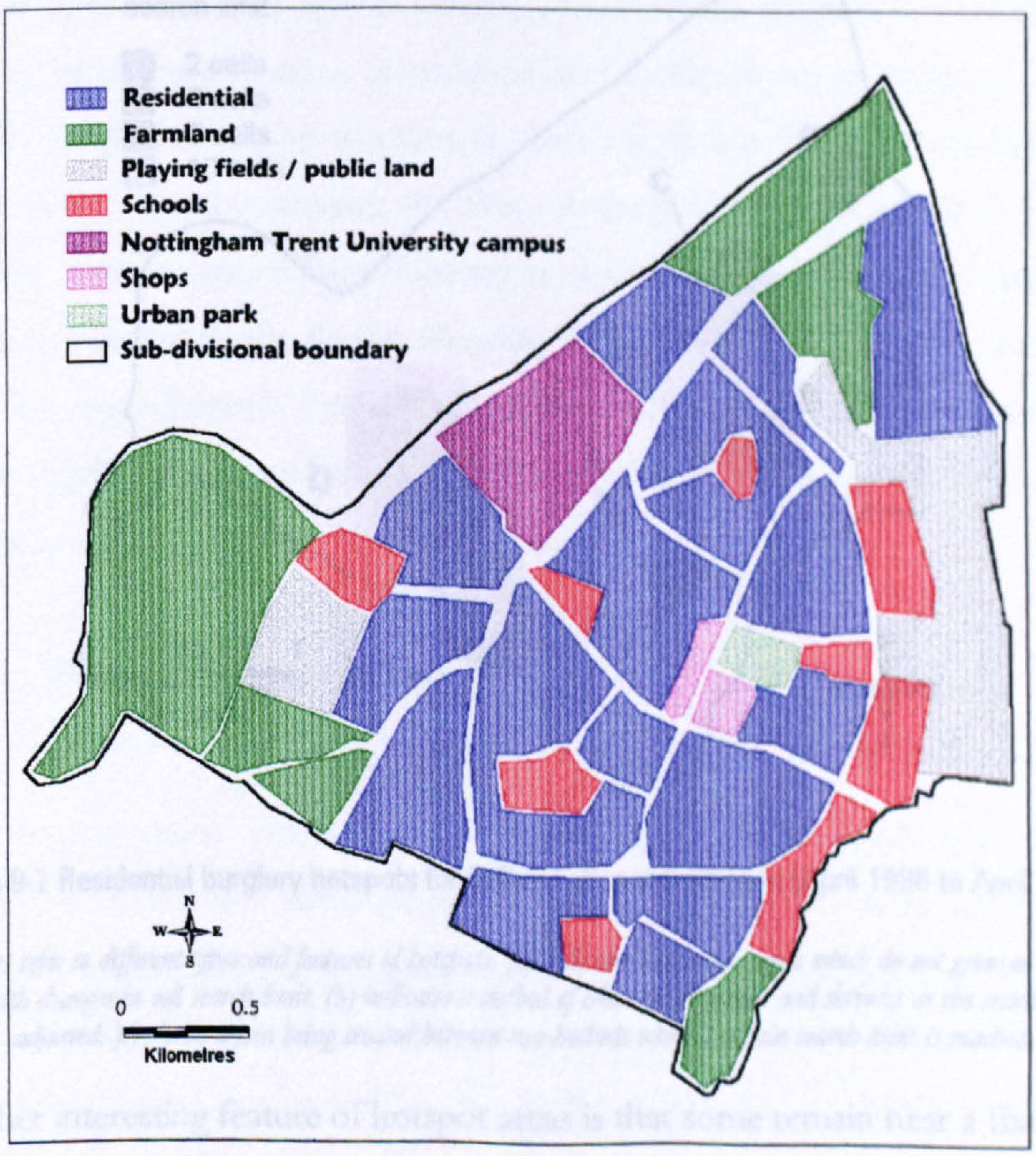


Figure 9-6 Simplified landuse map of Clifton.

There are a number of interesting features apparent in the hotspots generated by the process described in this chapter. The graphs in Figure 9-1 (on page 261) show the number of crimes per cell for each potential search limit. It is also possible to calculate a value for crimes per cell in each individual hotspot for a predetermined search limit. In this manner a hierarchy of hotspot intensity can be constructed and each of the hotspot regions can be graded on a scale of importance. The police and crime prevention authorities can then be made aware of both the statistically significant crime hotspots in the local area, and the intensity of each of the crime sites. An ability to determine a strategic ranking of hotspots permits the allocation of crime prevention resources in a logical and prioritised way and can lend scientific support to the process of resource allocation.

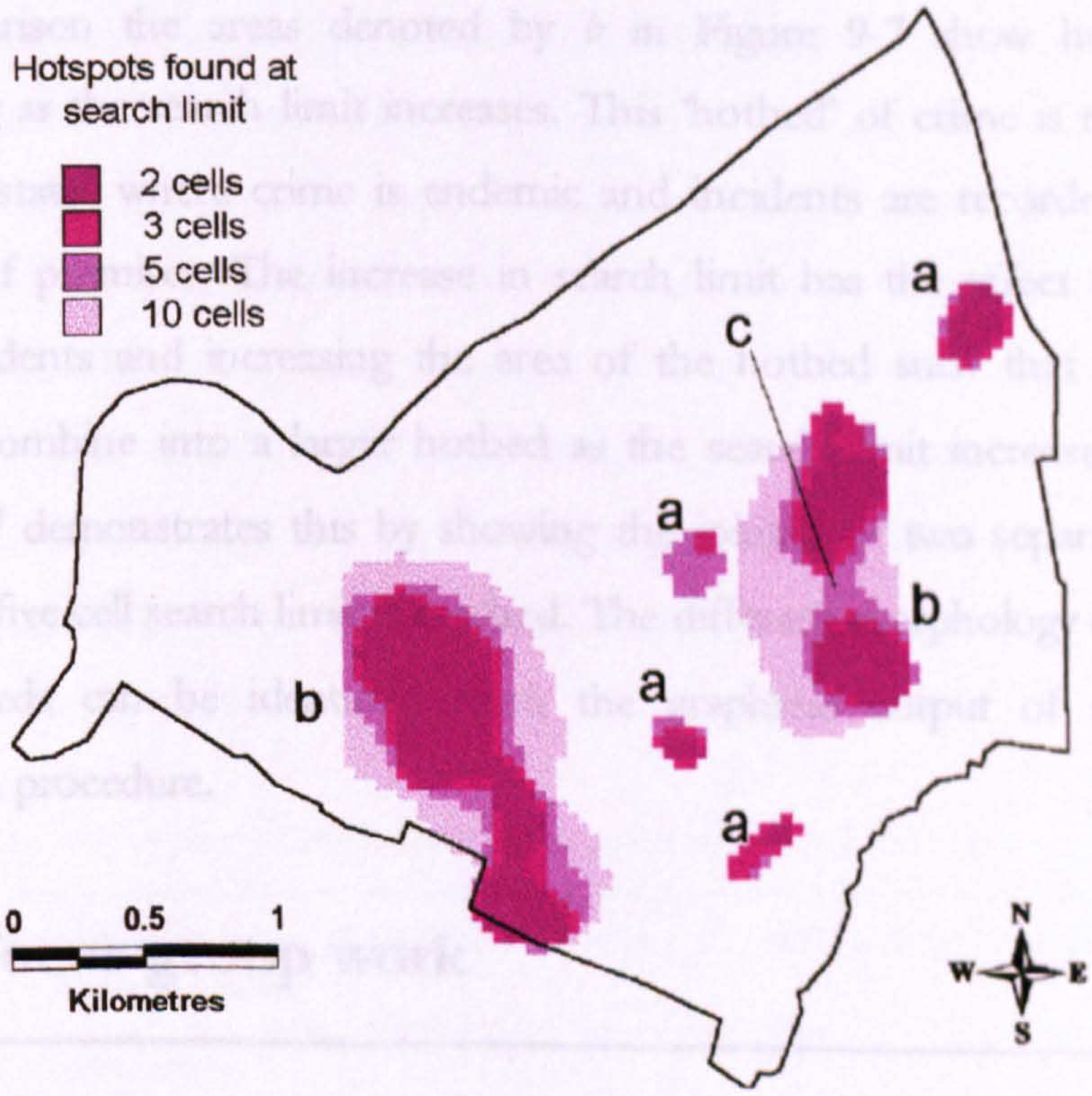


Figure 9-7 Residential burglary hotspots for Clifton police sub-division (April 1996 to April 1997).

Letters refer to different types and features of hotspots. (a) indicates hotpoints, areas which do not grow or shrink much with changes in cell search limit. (b) indicates a hotbed of crime which grows and shrinks as the search limit is adjusted.)c) shows a join being created between two hotbeds when a certain search limit is reached.

Another interesting feature of hotspot areas is that some remain near a fixed size (hotpoints) growing little as the search limit for the LISA statistic is increased, while others expand (hotbeds), change shape, and can coagulate with neighbouring hotspots. These effects can be seen particularly well in the Clifton sub-division residential burglary data shown in Figure 9-7. The hotspots identified as *a* are hotpoints which were often found to be caused by repeat victimisation at the same location. This has the effect of increasing the number of 'hits' at one single location. The lack of criminal activity in the surrounding area causes the area of the 'hotpoint' to remain stationary as the search limit increases. The lack of events in the vicinity can be linked to a number of reasons. It might be that the premises are very large (such as a block of flats) and all crimes recorded at that location are geocoded with the same grid reference. It is also possible that one location is particularly vulnerable to crime and is a magnet to criminal activity. Either of these reasons would generate a large number of incidents at one location to the exclusion of surrounding areas.

In comparison the areas denoted by *b* in Figure 9-7 show hotspot areas expanding as the search limit increases. This 'hotbed' of crime is noticeable in housing estates where crime is endemic and incidents are recorded at a large number of premises. The increase in search limit has the effect of including more incidents and increasing the area of the hotbed such that two smaller hotbeds combine into a larger hotbed as the search limit increases. Area *c* in Figure 9-7 demonstrates this by showing the joining of two separate hotspots when the five cell search limit is reached. The different morphology of hotpoints and hotbeds can be identified from the graphical output of the hotspot generation procedure.

9.4.1. Focus group work

To interpret more accurately the results from the hotspot analysis and perception study, three focus groups – one at each station – were set up. This was done to improve the interpretation of the results. These sessions were run with *mini-groups*, a term used to indicate groups sizes of 4-6 people, with the role of *moderator* being performed by the author (Greenbaum, 1993). Participant recruitment was left in the hands of the section inspectors due to their better knowledge of the available participants. The inspectors were asked to identify individuals who had an extensive understanding and experience of the crime problems of the area, and who would take an active interest in this type of research.

SESSION FORMAT

Each session had the same format in line with general focus group guidelines (Greenbaum, 1993 p.43). There was an initial *warm-up session* where the moderator was introduced (the participants knew each other already), and the participants were introduced to the moderator. A *details section* followed which discussed the technical aspects of the survey including; the mapping technique, the survey methodology and the display of the results. The important *key contents section* was next, which gave the respondents a wide rein to interpret the results of the survey and explain possible causes for some of the findings. The session was then directed by the moderator into a discussion of the means by which police officers receive information about crime and incidents on their area, and a discussion of crime in their area in general. This was the longest section of each focus group. The group was wrapped up with a *summary section* where the participants were asked for any last thoughts, the moderator summarised his notes, and checked with the group that his interpretation of the session was correct. A summary of the research objectives for the focus group part of the study in shown in Table 9-4.

Table 9-4 Statement of focus group research objectives.

The general objective was to interpret the results from the hotspot analysis and survey conducted on Trent division.
More specific questions:
Did the participants feel that crime distribution had changed significantly from last year to this year? (this question was aimed at addressing the mismatch in survey and crime data dates).
Did the participants feel that the hotspots indicated were an accurate representation of crime concentrations on their area?
Could the participants explain why autocrime survey results appeared poorer than the burglary results?

The focus groups were conducted in the briefing rooms of Clifton and the Meadows police stations, and in the conference room of West Bridgford station. Table 9-5 outlines the focus group details.

Table 9-5 Mini-group details of the police focus groups.

Station	Meadows	Clifton	West Bridgford
Date	11th May 1998	15th April 1998	15th April 1998
Time	9.30am – 10.20am	10.55am – 11.45am	12.15pm – 1.45pm
Location	Briefing room, Meadows police station	Briefing room, Clifton police station	Conference room, West Bridgford station.
Participants	1 x Inspector	1 x Inspector	1 x Inspector
	1 x Sergeant	1 x Sergeant	3 x PCs
	3 x PCs	3 x PCs	
Participants selected by	Inspector David Shardlow	Inspector Nigel Hallam	Inspector David Powell
	(sub-divisional commander)	(sub-divisional commander)	(sub-divisional commander)

FOCUS GROUP INTERPRETATION OF THE RESULTS

Clifton officers felt that the temporal discrepancy between the crime data and the survey dates was not a problem. The main residential burglary area is the Nobel Road area, and the adjoining private estate. The Nobel Road estate was built 20 years ago and stocked with problem residents from around the city; including residents from Hyson Green and Radford. There are apparently 7 or 8 problem families (and extended families) responsible for most, if not all, of the burglary crime. Examples from the officers included the case of one woman who was burgled 5 times by her next-door 'neighbours'. She finally moved out and a relative of the burglars moved in. Non-residential burglaries are focused on the local shops, council offices and library. The Nottingham Trent University campus in Clifton was highlighted as a hotspot by the computer process but had no officer-generated predictive crosses at that location. It was felt that any problems there were dealt with by the campus security "for what they are worth". Motor vehicle crime was also seen as being committed by youths from the Nobel Road area, who were not yet old enough to progress onto burglary. One problem which does not appear in the crime figures are the instances of stolen cars being found in the Nobel Road estate; cars that are stolen from elsewhere. This was seen as the reason for so many crosses in the area for motor vehicle crime.

West Bridgford officers also felt that the temporal discrepancy between the crime data and the survey dates was not a problem. Residential burglaries in the area are often caused by youths coming over the River Trent from Sneinton, St. Anne's and the Meadows area. They use either the road bridge, or if they are younger and have no vehicle access, the footbridges. The non-residential burglaries are located in expected areas, as shops and leisure centres are focused in quite specific places. This knowledge does not explain the poor estimates of non-residential burglaries (46.4%). There are now fewer burglaries at schools in the area because the schools have installed large security fencing around their premises. The motor vehicle crime on the sub-division is concentrated in specific car parks. There was a perceived misconception that autocrime was concentrated in the LadyBay area. The focus group participants felt that this may be because the LadyBay area is seen as a high crime area, irrespective of the crime type. Officers were choosing this area for their vehicle crime estimates because it was a busy area for burglaries and other incidents.

The Meadow officers identified the main burglary hotspots as the work of a small number of known individuals. This is reflected in the high success rate of their hotspots estimates (over 90%). One individual in question had now left the area and was not (at the focus group time) seen as a cause of any current burglary figures. The central group of burglaries occurred in the student accommodation of the old Meadows and were attributed by the officers to drug related crime. The high success rate for vehicle crime (60%) was attributed to a knowledge of the main activity near a local football ground. The primary vehicle crime hotspot was identified as the main Nottingham train station and this area is the preserve of the British Transport Police (BTP). Few officers chose this location for a hotspot estimate. The focus group felt that they would not expect to know about incidents happening there as Nottinghamshire Police and BTP do not share jurisdictions or crime data. This was even in light of the fact that the hotspot was generated from Nottinghamshire police crime data, and *not* British Transport Police data.

9.4.2. Seeing through the results

The initial interpretation of the hotspots generated by the SPAM program, and the estimates from the survey of officers identified that the computer calculated hotspots are located correctly and that there are definite reasons why particular hotspots were or were not detected.

The focus group discussions went further than this, and examined the underlying causes of the results, and in particular the low success rate of police identification of vehicle crime hotspots in relation to the residential burglary rate. Both of these crimes are high local priorities and yet the success rates in hotspot identification for vehicle crime never grew above 60%, and the success rates for residential burglary never dropped below 60%.

Before looking at the hotspot printout, Clifton officers felt that vehicle crime was sporadic and "all over the ground". Little enthusiasm was evident for tackling vehicle crime, and the emphasis of the officers was always placed on burglary. According to the officers the divisional intelligence unit rarely gives them intelligence targets to watch and when it does they are always burglary targets. Vehicle crime suspects were ignored as there are always "*better targets we can do something about*". Within the division there is an importance attached to burglary which is not the same for motor vehicle crime. To quote one PC: "*Vehicle crime is seen as an inevitability*".

The officers mentioned a previously operated system of placing crime locations on a map with pins. The usual problems with this type of crime mapping were cited as the reasons for its eventual demise (see chapter 2). It was noted that one system ground to a halt once the individual who was enthusiastic about the map was transferred to another station.

At West Bridgford, crime information comes from the briefing computer, a local system for passing information from one shift to another. Few officers attend vehicle crime calls (not just at West Bridgford, but across the division), and now with the division wide introduction of ***crime scene visitors***, fewer officers attend burglaries. Crime scene visitors are police officers dedicated to the reporting of burglaries with special training in crime prevention, the needs of the crime scene officer (forensics) and with a good knowledge of the local burglars. These officers are the first allocation choice when burglaries are to be reported. The focus group officers felt that if a burglary was attended and reported by a normal shift officer then it would appear on the computer and word of mouth might also spread details about the crime. This is now less applicable because officers do not attend crimes as often. The briefing computer also only highlights incidents over the last 24 hours.

Motor vehicle crime is not seen as important as burglary, even though it is a divisional and section objective. It was felt that motor vehicle crime was "all over the place" (even though the hotspot maps show this is not the case) and if the locations were read out on parade there was too much information. As the one PC said: "*A few locations you can remember but with autocrime you get swamped with information overload.*"

There was also the feeling that motor vehicle crime was becoming "*decriminalised*". Officers rarely attend the crime scene, and it was felt that crime prevention advice is rarely distributed. The officers thought that the victim is generally only interested in obtaining a crime report number which is needed for the insurance claim. To quote one officer: "*We do not consider autocrime to be nearly as important as burglary.*"

At the Meadows station, information is passed from shift to shift by writing the information on a whiteboard. This was felt to be an acceptable means of information dissemination because the station area is small and the numbers of incidents are relatively few. The burglary and vehicle crime locations are written onto the board and are available for all to see, though the Sergeant did say that "*You can come on in the morning and find 20 autocrimes and only 2 burglaries which are much easier to remember*". Vehicle crime was again seen as a low priority. The crime desk deal with most of the vehicle crime. Only if two specific criteria are met will an officer attend a vehicle crime incident: if there is the presence or knowledge of a suspect, or the possibility of fingerprint evidence.

Further quotes include comments like: "*Well, nobody gets hurt or that upset by it [motor vehicle crime], not like they do at a burglary*". With regard to burglaries, the inspector felt that the use of Crime Scene Visitors meant that they "*take away the knowledge base of the officers*" regarding burglaries.

Summary

In summary, the focus groups produced similar results at all three stations.

- Burglary is seen as a higher priority than vehicle crime.
- Burglary is seen as a preventable crime whereas vehicle crime is seen as an inevitability.
- The use of burglary scene visitors to report burglaries was perceived to reduce the local crime knowledge of officers.
- Vehicle crime is perceived to be endemic to all areas of the division, though this is not the case when the hotspots are viewed in Figure 9-2, Figure 9-3, and Figure 9-4.
- Vehicle crime is seen as a lost cause and neither the police nor the public are interested.
- Because vehicle crimes are rarely reported their locations are not known.
- The current means for disseminating the vehicle crime location information are not working (Meadows was the exception) as there is either too much information or too little.

9.4.3. Implications of this work for policing policy

This survey has raised a number of issues which are relevant to current policing policy. There are two aspects to the work which are important: the accurate identification of problems (hotspots), and the passing of that information on to the patrolling officers who are responsible for combating local crime. The first issue of correctly identifying problems has come to the fore with the introduction of the Crime and Disorder Bill 1998. Clauses 6 and 7 of this act of government require the police to act in conjunction with local authorities in drawing up a local action plan to combat crime problems, the first stage of which is the publishing of a local crime audit by April 1999 (Home Office, 1998). The importance of accurately identifying significant hotspots of crime can therefore be seen as important in the context of this government bill.

The second aspect of the work is the dissemination of this information to those charged with the responsibility for crime prevention. Although the police have a relatively minor role in the detection of criminal activity and less than 20% of recorded crime comes to notice directly through the actions of the police (Bottomley and Coleman, 1976), it should be a priority to improve the patrolling patterns of the police in order to get them in the right place and the right time, *more of the time.* It is also impossible to calculate the amount of successful crime prevention which takes place by the simple act of a police officer or car patrolling down a road and deterring an imminent criminal act.

This study has shown that while officers have a good knowledge of the burglary patterns, they are less successful with vehicle crime. The current dissemination methods of whiteboards and text on computers is generally not successful in imparting the areas of concern to officers.

9.5. CONCLUSIONS

This chapter has advanced the hotspot analysis from the previous chapter and demonstrated a use with a practical application. The selection of a suitable bandwidth for search circle size is an important part of the application process. The US STAC system suggests a bandwidth of 220 metres, but this is based on the street layout of Chicago. The analysis presented here suggests that a survey of the officers is the most effective means of calculating the extent of likely hotspots by asking the police officers themselves. The range of chosen radii values ranged between 100 metres and 400 metres. It is noticeable that the Chicago suggestion falls within this band.

A minor modification to the Getis and Ord G_i^* statistic is suggested to remove the effect created by large blank areas of no crime in the study area, and this adds to the power of the statistic by including only those areas where crime is committed. The choice of a suitable search limit for the LISA statistic is also important. "No matter what the statistic, a key question that must be addressed in these operations is the definition of what is meant by local." (Unwin, 1996 p.547) Cross-comparison of hotspot areas and crimes per cell indicates a cell range of between 4 and 5, which in the examples chosen closely matched the original bandwidth for the search circle.

This combined system has been applied to three sub-divisions of Nottinghamshire Constabulary. It was noticeable that two additional types of hotspot were evident when the results from different search limits were combined. Figure 9-7 shows these hotpoints and hotbeds created by two different processes. The hotpoints are the results of a succession of incidents at a single location remote from nearby crimes while the hotbeds are generated by a many randomly scattered crimes at different locations but near to each other.

Focus mini-groups were used in an attempt to further analyse the results of the hotspot analysis when compared with a police perception survey of hotspot location. The results of the focus group work show that information about vehicle crime incidents is not filtering down to the officers responsible for the prevention of these crimes. It is impossible to tell if a crime mapping system could improve the situation, though it can be seen that the current intelligence dissemination systems in the stations are not entirely successful. These results could be interpreted as justification for trying a new graphical system. An identical repeat survey conducted six months or a year after the introduction of such a system could assess its impact.

The question was asked at the start of this chapter as to whether a police officer's impression of the extent of criminality on his beat is accurate. This chapter examined the hypothesis that police officers could construct an accurate perception of crime distribution from exposure to daily policing practices by comparing crime hotspots generated from the recorded crime data with the perceived local hotspots catalogued from surveys with police officers. It would appear that this is the case for burglary crimes because the officers are interested in its prevention and detection, and there are not so many incidents that officers get 'information overload'. They are much less successful with vehicle crime. This survey and analytical process could be used both as a reason and a method for introducing a mapping system designed to improve the awareness of vehicle crime incidents.

10. Conclusion

This chapter draws the thesis to a conclusion, reviews the thesis in relation to the aims stated in the introduction and identifies directions for future research in the area of spatial and temporal crime analysis.

10.1. SUMMARY

In the introduction to this thesis (Chapter 1) the stated intention was to investigate areas of concern in current crime management, and to examine where new spatial and temporal investigation techniques could improve the analysis of high volume crime at a local community level. Until recently in most police services, the only spatial and temporal analysis of crime was conducted by statisticians at the force headquarters with little or no regard for any short term or localised patterns of crime. The recent move within British policing towards a more decentralised, proactive style has shifted the analytical focus onto analysts and intelligence officers at the police divisional level. The intelligence officer at a divisional station is now expected to be the hub of the local intelligence gathering effort. However, for high volume crime, this has left an analytical void. Force level analysis techniques are neither appropriate nor subtle enough to elicit any meaningful information at a local level from the mass of crime data generated within the police service.

There was a broad range of possible directions the study could take. It was decided to focus on the immediate needs of the police service at a local level and in particular on the need for improved identification of repeat victimisation, and on more accurate methods of hotspot analysis. These requirements are reflected in the level of funding available (£32 million initially from a £250m budget) for the current Home Office targeted policing initiative.

Specific conclusions have been drawn in a summary at the end of each chapter. A brief outline of these findings is presented below.

Any analysis of crime must be placed within the correct temporal context. Therefore the first two analytical chapters of this thesis advanced both a new method of selecting accurately temporally unspecific crimes within a search criteria, and identified distinct spatial and temporal variations in assault and disorder criminal activity in Mansfield. The use of aoristic and probabilistic aoristic search techniques (described in Chapter 4) corrects errors in temporal search techniques that occur when more simplistic search algorithms are used with non-specific crime data. Time series analysis techniques are used in Chapter 5 to identify repetitive patterns in assaults and disorder activity. This chapter has identified important geographical and temporal variations between crimes and incidents.

The use of georeferenced crime data analysed within a GIS has been proposed in this thesis as an improved method of identifying repeat victimisation (Chapter 6: Identifying repeat victimisation), and a spatial analysis using this technique in Chapter 7 has revealed new aspects to the distribution of burglary crime with respect to local deprivation levels. The use of GIS and georeferenced crime data has been shown to dramatically improve the rapid detection of burglary repeat victimisation. The development of a new Vicinity-based aggregation method has been successful in preserving data integrity and avoiding minor crime location georeferencing errors, and it has identified an important pattern in the distribution of repeat victimisation locations in relation to the deprivation index of an area. The recognition that repeat victimisation is concentrated in more deprived areas will be of benefit to crime prevention agencies.

The development of hotspot analysis techniques is a growth area within crime analysis and is seen as one of the positive ways in which spatial analysis can usefully aggregate high volume crime data. Numerous techniques are available either as freeware or commercially, though these techniques have limitations as shown in Chapter 8 (Hotspot analysis). A two stage process using a localised surface generation algorithm followed by a Local Indicator of Spatial Association (LISA) statistic is proposed in Chapter 8 as the best method available for detecting crime hotspots. It provides a hotspot detection program with a statistical confidence limit. It is to be hoped that dissemination of this technique may help to promote more advanced algorithms and processes into the field of crime mapping. The ubiquity of inferior methods can be seen with the dominance of STAC in a number of recent books (for example: LaVigne and Wartell, 1998; Weisburd and McEwen, 1998).

This thesis reveals patterns in community level crime which have not been recognised previously as traditional techniques in spatial and temporal investigation have lacked the analytical ability. A number of new techniques are presented which are geared towards the needs of a crime analyst at a divisional police station, an individual who has until now lacked the necessary analytical tools to perform the role effectively.

The current changes in British policing that are being mirrored elsewhere around the world are changing the way that intelligence is gathered within the police service. This is likely to see continued development of local high volume analysis techniques as new possibilities are considered by police officers and academics. It is also likely that a revision of the place of crime mapping and analysis within the integrated crime management model (Amey *et al.*, 1996) will prompt changes in focus and drive new techniques.

10.1.1. Further work

Full many a flower is born to blush unseen,
And waste its sweetness on the desert air. [1]

Thomas Gray's words are perhaps indicative of the fate of theory and research without a practical application. It is essential to test the theories and applications presented here in a real world scenario. It was the intention of this research to include the development and testing of a mapping and analysis system in a number of stations of Nottinghamshire Constabulary. However over a four year period a number of barriers were placed in the path of this aim and a degree of intransigence within certain departments of the force prevented implementation of a freely-offered system. This was unfortunate, as the development of such a system would have shown how well the theoretical side could have been implemented.

1 Elegy Written in a Country Churchyard (1750), Thomas Gray, 1716-1771

This thesis is not a panacea to all high volume crime problems as the subject is too vast and complex to attempt such a task. It is designed to be a contribution of ideas and techniques towards a greater understanding of the spatial and temporal aspects of crime at a particular level. The research presented here has shown that large scale and long term analytical techniques are less appropriate at a divisional or community level and that specific methods of analysing crime and incidents are necessary if significant and subtle local patterns of criminal behaviour are to be recognised. It has been shown that there are specific theoretical considerations which have to be, and have been, answered. The next stage is to put this theory into practice.

The logical next step would be the development of a functional mapping system inside a police service. The creation of a productive system is necessary to test the theories and practices which have been developed in this research, and it is imperative to implement a working system from the mosaic of theoretical techniques presented both here and in the work of other researchers in the field, to see if the theoretical ideas presented in the research are valid and will work in the real world.

An additional benefit of a production system is that an existing system promotes further research in particular areas. A lack of technical knowledge within the police service can sometimes lead to a lack of imagination as to the possibilities that a system can offer. The existence of a working system could stimulate interest in new developments. A potential developmental process could occur in three stages.

Phase One Initial development of a basic system to stimulate interest and provide basic functionality. This type of system would include basic crime selection queries, a number of pre-programmed search queries for commonly-performed tasks, a simplified 'press and print' facility which added automatically the force logo and confidentiality warnings to the hardcopy output, and a simplified menu facility to enable usage by non-GIS experts. The purpose of such a system would be to achieve acceptance within the force. It is possible that this phase could be achieved within a period of 3 months.

Phase Two A more advanced second stage system would the full analytical processes described in this thesis, including an advanced repeat victimisation identification option, improved temporal search facilities, aoristic histograms, graphical output, and automated linkage with an advanced hotspot detection process. The purpose of this stage would be to show the possibilities of what could be achieved. A development time of one year is not unreasonable for this phase of a system.

Phase Three This stage would include the practical application of on-going research activities and could initiate and drive research in locational crime analysis into the next decade.

It would be hoped that a phase three system would focus and drive the research work by identifying the potential future areas of research and development.

To achieve any of these stages, an initial (stage one) development application must be in place to drive understanding of further development. New areas of research after this could include:

- Examining the geographical relationship between the location of victim and the location of suspects,
- The introduction of suspect identification to assist with detection of other nearby incidents,
- The development of more detailed MO (*modus operandi*) analysis,
- A recognised process of identifying linked incidents and crimes,
- The impact of hotspots on policing strategy,
- The implementation of crime mapping within an integrated crime management model.

10.1.2. Beyond crime

In conclusion, it should be recognised that the processes and techniques presented in this thesis may have applications beyond the fields of policing and crime prevention. This final section describes where each of the main thesis research areas may have application beyond the crime-specific areas described in the previous chapters.

The aoristic temporal analysis process is a new yet conceptually simple approach to the analysis of temporally unspecific data. The ideas and analytical technique presented here are a 'first pass' at a methodology and it must be understood that the process is likely to undergo a number of refinements and revisions before a definitive framework is complete. Types of refinement could include a greater understanding of the impact of temporal scale. The analysis presented in chapter 4 looked at the aoristic values on a day-to-day analysis, but a comprehension of the affect of an hourly or twice hourly analysis would be beneficial.

This type of work has application beyond the limitations of the crime field here and is relevant to many types of spatial work that contain a continuous temporal element. Langran (1989) recognised the need for an understanding of this fourth temporal dimension in spatial studies lamenting the lack of frameworks that could cope with temporally continuous data. It is hoped the aoristic aspect of the thesis is a contribution in this direction.

Chapter 6 (Burglary, victimisation and deprivation) introduced a technique for estimating a value for an area in the immediate vicinity of a point from the values of surrounding polygons. Although the application presented in this thesis referred to a crime or repeat victimisation location, the process is equally applicable whenever the extraction of an local variable surrounding a point is required. This type of process could be applied in situations where there is some doubt as to the accuracy of the point location. One area of application includes the field of epidemiology where a researcher might be interested in estimating the socio-economic factors in the area in the immediate vicinity of the victim of a rare disease. The search for contributory factors in the development of rare diseases can often include an analysis of the immediate environmental and socio-economic conditions of the victims location. This process would allow a researcher to estimate a value for each victims location with more certainty in the knowledge that any errors in the effect of spatial inaccuracy of the victim's location will be dramatically reduced by the local spatial weighting process that takes place within the analysis. This allows the researcher to maintain the local focus of the experiment whilst increasing confidence in the final result.

This process could also be applicable in other countries that lack the resolution of the Ordnance Survey AddressPoint data set and have to rely on geocoding products with a lower spatial precision.

Another area of application is in the understanding of the spatial variation in wildlife habitats. On rare occasions, the accuracy of the habitat map can exceed the accuracy of the wildlife field survey. This can occur when vegetation cover is accurately digitised from aerial photography or ground survey and the wildlife survey, such as a bird location survey, is generated from visual sightings recorded on a less accurate local map (Lavers *et al.*, in lit). In this situation, the accuracy of the wildlife locations point data file is an essential factor in the result of the experiment, and a process such as employed by the Vicinity program can reduce the potential errors and allow for vagaries in the veracity of the point data file.

The hotspot techniques presented in Chapters 8 and 9 describe a variety of different approaches to the problem of identifying clusters in point data sets. These techniques have merit in any field that considers the lack of randomness in the density and spatial distribution of discrete locations. The techniques that utilise SPAM and the further addition of the Getis and Ord G_i^* local statistical process will be of benefit to many practitioners wishing to include a rigorous statistical element to their spatial analysis. Development of the programming to simplify the process for GIS users and take the process beyond the realm of developers will be a necessary piece of further work. By doing so the technique will become more accessible to less experienced GIS users and add a statistical element to the spatial analysis work they perform that has until now been unavailable without considerable programming expertise.

The most obvious application area for this type of local spatial statistics beyond crime prevention and policing is again in the field of epidemiology. Much of the pioneering original work into the search for discrete clusters in spatial point data sets can be found in the literature surrounding the search for clusters in rare diseases (the literature is thoroughly reviewed by Gatrell and Dunn, 1995).

Epidemiology differs in application from police work in one key factor. A police officer is concerned with absolute hotspots of crime, that is to say that even if there are no people living in a hotspot area (such as an industrial estate), the officer would still be concerned about the level of crime. If an epidemiologist adopted the same approach then they would discover that most cases of rare diseases occur in the large metropolitan areas that have the highest population densities. This ecological fallacy has lead to an appreciation in the medical field that the distribution of the population in an essential key in understanding the relative incidence rate of the disease in question.

The difference between the desire to detect an absolute statistical concentration of points in the data set, and to detect a relative cluster of points in the data set taking into consideration the underlying population density is the key to understanding the fundamental differences in analytical approach.

The dual technique of a SPAM application followed by the use of the Getis and Ord G_i^* local statistical process is a unique methodology for the detection of local clusters in point data files, the Getis and Ord G_i^* statistic only having been applied previously to data in polygon format. A degree of reprogramming to appreciate the underlying population geography would be necessary before the whole process would be of use to epidemiologists, though this is not an unfathomable task and would open to epidemiologists the new branch of LISA statistics.

Many of the techniques presented in this thesis are discrete techniques that could and perhaps should be applied to a variety of data sets. This would enable both the checking of the veracity of the processes across different data structures, and may suggest further areas of application and further areas in the development of the methodologies concerned. It is to be hoped that development of these processes will advance the techniques into areas beyond the limited scope of crime prevention and policing.

References

Alder, H.L. and Roessler, E.B. (1962), *Introduction to probability and statistics.* San Francisco, London: Freeman.

Allen, R. (1996), 'Socioeconomic conditions and property crime - a comprehensive review and test of the professional literature', *American Journal of Economics and Sociology,* 55/3: 293-308.

Amann, R. (1996), 'Keynote speech to ACPO', *ACPO Conference,* London, ESRC.

Amey, P., Hale, C. and Uglow, S. (1996), 'Development and evaluation of a Crime Management Model', *Police Research Group: Police Research Series,* Paper 18: 37.

Anderson, D., Chenery, S. and Pease, K. (1995), 'Biting back: Tackling repeat burglary and car crime', *Police Research Group: Crime Detection and Prevention Series,* Paper 58: 57.

Anselin, L. (1995), 'Local Indicators of Spatial Association - LISA', *Geographical Analysis,* 27/2: 93-115.

Audit Commission (1996), 'Streetwise: Effective Police Patrol'. London, Audit Commission.

Bailey, T.C. and Gatrell, A.C. (1995), *Interactive Spatial Data Analysis.* London: Longman.

Baldwin, J. and Bottoms, A. (1976), *The Urban Criminal.* London: Tavistock.

B.C.M.P. (1997), 'Crime Awareness News'. Brent Crime Mapping Project, London: 4.

B.C.M.P. (1998), 'Crime Awareness News'. Brent Crime Mapping Project, London: 4.

Berkeley Police Department (1997), 'The Police Department GIS Prototype', http://www.ced.berkeley.edu/aegis/oma/orhtml.html.

Berry, W. and Jones, H. (1995), 'Investigating spatial patterns of crime through police statistics, crime surveys and census profiles - findings from Dundee housing estates', *Scottish Geographical Magazine,* 111/2: 76-82.

Besag, J. and Newell, J. (1991), 'The detection of clusters in rare diseases', *Journal of the Royal Statistical Society, A,* 154/1: 143-155.

Block, C.R. (1984), 'How to handle seasonality'. Chicago, Illinois Criminal Justice Information Authority.

Bottomley, A. and Coleman, C.A. (1976), 'Criminal statistics: The police role in the discovery and detection of crime', *International Journal of Criminology and Penology,* 4: 33-58.

Bradbury, J. (1981), 'An investigation of the spatial distribution of crime in Greater Nottingham', Geography, University of Nottingham.

Brantingham, P.J. and Brantingham, P.L. (1981), 'Notes on the Geometry of Crime', in P. J. Brantingham and P. L. Brantingham, eds., *Environmental Criminology*. 27-54. Sage: London.

Brown, M. (1982), 'Modelling the spatial distribution of suburban crime', *Economic Geography*, 58: 247-261.

Burquest, R., Farrell, G. and Pease, K. (1992), 'Lessons from schools', *Policing*, 8: 148-155.

Campbell, G. (1992), 'GIS in the police environment', in J. Cadoux-Hudson and I. Heywood, eds., *Geographical Information 1992/3: The Yearbook of the AGI*. 114-118. Taylor & Francis: London.

Cater, J. and Jones, T. (1992), 'Crime and disorder', in Cater, J. and Jones, T. eds., *Social geography: an introduction to contemporary issues*. 79-113. Edward Arnold: London.

Chatfield, C. (1989), *The analysis of time series: An introduction*. London: Chapman and Hall.

Chenery, S., Holt, J. and Pease, K. (1997), 'Biting Back II: Reducing repeat victimisation in Huddersfield', *Police Research Group: Crime Detection and Prevention Series*, Paper 82: 41.

Clegg, M. and Robson, M. (1995), 'May the force be with you: as quickly as possible using GIS', Mapping Awareness, 9: 30-33.

Cohen, L.E. and Felson, M. (1979), 'Social change and crime rate trends: A Routine Activity Approach', *American Sociological Review*, 44: 588-608.

Coleman, A. (1985), *Utopia on trial: vision and reality in planned housing*. London: Hilary Shipman.

Coleman, C. and Moynihan, J. (1996), *Understanding crime data*. Buckingham, UK: Open University Press.

Crime Concern (1996), 'Mansfield Residential Crime Survey'. Mansfield, Nottinghamshire, Crime Concern.

Criminal statistics (1997), 'Criminal statistics: England and Wales 1996'. HMSO, London.

Davis, J.C. (1986), *Statistics and data analysis in geology*. London: John Wiley.

Diez, L. (1995), 'The use of call grading: How calls to the police are graded and resourced', *Police Research Group: Police Research Series*, Paper 13: 46.

Dorling, D. and Openshaw, S. (1992), 'Using computer animation to visualize space-time patterns', *Environment and Planning B - Planning & Design*, 19/6: 639-650.

Ebdon, D. (1996), *Statistics in Geography*. Oxford: Blackwell.

Economist (1994), 'Crime prevention: Security forces', The Economist, 334: 30.

Ekblom, P. (1988), 'Getting the best out of crime analysis', *Police Research Group: Crime Prevention Unit Series*, Paper 10: 1-42.

Ellingworth, D., Farrell, G. and Pease, K. (1995), 'A victim is a victim is a victim?', *British Journal of Criminology*, 35/3: 360-365.

Elliott, C. and Ellingworth, D. (1996), 'The relationship between unemployment and crime - a cross-sectional analysis employing the British Crime Survey 1992', *International Journal of Manpower*, 17/6-7: 81.

Environment, The department of (1995), 'Index of local conditions'. London, Department of the Environment.

Farrell, G. and Pease, K. (1993), 'Once bitten, twice bitten: Repeat victimisation and its implications for crime prevention', *Police Research Group: Crime Prevention Unit Series*, Paper 46: 32.

Farrington, D. (1996), 'Quantitative criminology in the UK in the 1990's - a brief overview', *Journal of Quantitative Criminology*, 12/3: 249-263.

Farrington, D.P. and Dowds, E.A. (1985), 'Disentangling criminal behaviour and police reaction', in D. P. Farrington and J. Gunn, eds., *Reaction to crime: The public, the police, courts and prisons*. John Wiley: Chichester.

Field, S. (1992), 'The effect of temperature on crime', *British Journal of Criminology*, 32/3: 340-351.

Fox-Clinch, J. (1997), 'Crime and the digital dragnet', Mapping Awareness, 11: 22-23.

Gatrell, A.C. (1989), 'On the spatial representation and accuracy of address-based data in the UK', *International Journal of Geographical Information Systems*, 3/4: 335-348.

Gatrell, A.C., Bailey, T.C., Diggle, P.J. and Rowlingson, B.S. (1996), 'Spatial point pattern analysis and its application in geographical epidemiology', *Transactions, Institute of British Geographers*, NS 21: 256-274.

Gatrell, A.C. and Dunn, C.E. (1995), 'Geographical information systems and spatial epidemiology: Modelling the possible association between cancer of the larynx and incineration in North-East England', in M. J. C. DeLepper, H. J. Scholten and R. M. Stern, eds., *The added value of* *geographical information systems in public and environmental health.* 215-235. Kluwer academic publishers: London.

Georges, D.E. (1978), 'The geography of crime and violence: a spatial and ecological perspective', *Association of American Geographers: Resource papers for college geography*, 78/1.

Getis, A. and Ord, J.K. (1992), 'The analysis of spatial association by use of distance statistics', *Geographical Analysis*, 24/3: 189-206.

Getis, A. and Ord, J.K. (1996), 'Local spatial statistics: an overview', in P. Longley and M. Batty, eds., *Spatial Analysis: Modelling in a GIS environment.* 374. GeoInformation International: London.

Gottlieb, S., Arenberg, S. and Singh, R. (1994), *Crime Analysis from First Report to Final Arrest.* Alpha Publishing.

Graef, R. (1989), *Talking Blues: The police in their own words.* London: Collins Harvill.

Greenbaum, T.L. (1993), *The handbook for focus group research.* New York: Lexington.

Grescoe, T. (1996), 'The geography of crime', Geographical magazine (9): 26-27.

Grogger, J. and Weatherford, M. (1995), 'Crime, policing and the perception of neighbourhood safety', *Political Geography*, 14/6: 521-541.

Guerry, A.-M. (1833), *Essai sur la statistique morale de la France: precede d'un rapport a l'Academie de sciences.* Paris: Chez Crochard.

Hakim, C. (1982), 'The social consequences of high unemployment', *Journal of Social Policy*, 11: 433-467.

Harlequin (1997), 'Harlequin Criminal Intelligence System', http://webserver.harlequin.com/products/hcis/hcis.html.

Hatchard, G.D., Deans, A.J., Cavanaugh, K.L. and Orme-Johnson, D.W. (1996), 'The Maharishi effect - a model for social improvement - time-series analysis of a phase-transition to reduced crime in Merseyside Metropolitan area', *Psychology Crime and Law*, 2/3: 165-174.

Herbert, D.T. (1976), 'The study of delinquency areas: a social geographical approach', *Transactions, Institute of British Geographers*, NS1: 472-492.

Herbert, D.T. (1982), *The Geography of Urban Crime.* London: Longman.

Hirschfield, A. and Bowers, K.J. (1997), 'The effect of social cohesion on levels of recorded crime in disadvantaged areas', *Urban Studies*, 34/8: 1275-1295.

Hirschfield, A., Bowers, K.J. and Brown, P.J.B. (1995), 'Exploring relations between crime and disadvantage on Merseyside', *European Journal on Criminal Policy and Research*, 3/3: 93-112.

Hirschfield, A., Brown, P. and Todd, P. (1995a), 'GIS and the analysis of spatially-referenced crime data: Experiences in Merseyside, UK', *International Journal of Geographical Information Systems*, 9/2: 191-210.

Hirschfield, A., Yarwood, D. and Bowers, K. (1997), 'Crime Pattern Analysis, Spatial Targeting and GIS: The development of new approaches for use in evaluating Community Safety initiatives.', *Crime and health data analysis using GIS*, Sheffield, SCGISA.

HMIC (1997), 'Policing with intelligence'. London, Her Majesty's Inspectorate of Constabulary.

HMIC (1997a), Her Majesty's Inspectorate of Constabulary '1997.98 Primary inspection: Nottinghamshire Constabulary'. London, HMSO.

Home Office (1998), 'Crime and Disorder Bill'. London, House of Commons.

Home Office (1999), 'Targeted Policing: Prospectus for developmental projects'. London, http://www.homeoffice.gov.uk/cb102.htm.

Hope, T. (1995), 'The flux of victimization', *British Journal of Criminology*, 35/3: 327-342.

Hough, M. and Lewis, H. (1989), 'Counting crime and analysing risks: the British Crime Survey', in D. J. Evans and D. T. Herbert, eds., *The Geography of Crime*. Routledge: London.

i2 (1997), 'i2 Crime Analyst's notebook', http://www.i2ltd.demon.co.uk/.

ICJIA (1996), Illinois Criminal Justice Information Authority, *STAC user manual*. Chicago.

ICL (1995), 'Press release: ICL take command at Nottinghamshire Police', http://www.icl.com/news/press_releases/03May95.html.

Ireland, P. (1998), 'Helping police with their enquiries', Mapping Awareness, 12: 20-24.

Johnson, S.D., Bowers, K. and Hirschfield, A. (1997), 'New insights into the spatial and temporal distribution of repeat victimization', *British Journal of Criminology*, 37/2: 224-241.

Kendall, M.G. and Stuart, A. (1966), *The Advanced Theory of Statistics*. New York: Hafner Publishing.

Kennedy, L.W. and Veitch, D. (1997), 'Why are crime rates going down? A case study in Edmonton', *Canadian Journal of Criminology*, 39/1: 51-69.

Klinger, D. (1997), 'Negotiating order in patrol work: An ecological theory of police response to deviance', *Criminology*, 35/2: 277-306.

Knox, E.G. (1964), 'Epidemiology of childhood leukaemia in Northumberland and Durham.', *British Journal of Preventive and Social Medicine*, 18: 17-24.

Langran, G. (1989), 'A review of temporal database research and its use in GIS applications', *International Journal of Geographical Information Systems*, 3/3: 215-232.

Langran, G. (1993), 'Issues of implementing a spatiotemporal system', *International Journal of Geographical Information Systems*, 7/4: 305-314.

Lavers, C., McCullagh, M. and Fuller, R. (in lit.), 'A new method of investigating wildlife-habitat associations using aerial photography and GIS', *Geographical Systems*.

LaVigne, N. and Wartell, J. (1998), *Crime Mapping Case Studies: Successes in the field.* Washington DC: Police Executive Research Forum.

Lewton, P. (1998), 'Spatio-temporal crime pattern analysis', School of Geography. Nottingham, University of Nottingham: pp. 82.

Lowell, K.E. (1994), 'Probabilistic temporal GIS modeling involving more than 2 map classes', *International Journal of Geographical Information Systems*, 8/1: 73-93.

MacCleary, R. and Hay, R.A.J. (1980), *Applied time series analysis for the social sciences.* London: Sage Publications.

MacEachren, A. (1994), 'Time as a cartographic variable', in H. Hearnshaw and D. Unwin, eds., *Visualisation in Geographical Information Systems.* 115-130. John Wiley: London.

Maguire, D. (1991), 'An overview and definition of GIS', in D. J. Maguire, M. F. Goodchild and D. W. Rhind, eds., *Geographical Information Systems: Principles and applications.* 9-20. Longman: London.

Maguire, M. (1997), 'Crime statistics, patterns and trends: Changing perceptions and their implications', in M. Maguire, R. Morgan and R. Reiner, eds., *The Oxford Handbook of Criminology.* 135-188. Clarendon Press: Oxford.

Maltz, M.D., Gordon, A.C. and Friedman, W. (1991), *Mapping Crime in its Community Setting: Event Geography Analysis.* New York: Springer Verlag.

Mantel, N. (1967), 'The detection of disease clustering and a generalised regression approach.', *Cancer Research*, 27: 209-220.

MapInfo (1997), 'Mapplication: Crime Analysis', http://www.mapinfo.com.

Mayhew, H. (1862), *London Labour and the London Poor.* London: Griffin Bohn.

Mayhew, P., Maung, N. and Mirrless-Black, C. (1993), 'The 1992 British Crime Survey', *Home Office Research Study*, 132.

McCabe, S. and Sutcliffe, F. (1978), *Defining crime: A study of police decisions*. Oxford: Blackwell.

McCullagh, M.J. (1998), 'Quality, use and visualisation in terrain modelling', in S. N. Lane, K. S. Richards and J. H. Chandler, eds., *Landform monitoring, modelling and analysis*. Chapter 5, 95-117. Wiley: London.

Miller, H. (1991), 'Modelling accessibility using space-time prism concepts within GIS', *International Journal of Geographical Information Systems*, 5: 287-301.

Mirrlees-Black, C., Mayhew, P. and Percy, A. (1996), 'The 1996 British Crime Survey', *Home Office Statistical Bulletin*/19/96: 78.

Mitchell, D. (1997), 'Crime, policing and GIS: An emerging technology?', *Association for Geographic Information publication*: 10.

Morgan, R. and Newburn, T. (1997), *The future of policing*. Oxford: Oxford University Press.

Nagle, G. (1995), 'Urban crime: a geographical perspective', Geographical Magazine: 56-57.

Neave, H. (1978), *Statistical tables*. London: Unwin Hyman.

Newman, O. (1972), *Defensible space: Crime prevention through urban design*. New York: MacMillan.

Northwood Geoscience (1998), *Vertical Mapper version 2.0 manual*. Nepean, Ontario, Canada.

Nulph, D., Burka, J. and Mudd, A. (1997), 'Technical approach to developing a Spatial Crime Analysis System with ArcView GIS'. Washington, US Department of Justice.

Ohlemacher, T. (1995), 'An ecological regression-analysis of crime and poverty - an old tale told once again', *Kolner Zeitschrift fur Soziologie und Sozialpsychologie*, 47/4: 706.

Openshaw, S. (1984), 'The modifiable areal unit problem', *Concepts and Techniques in Modern Geography*, 38: 41.

Openshaw, S. and Charlton, M. (1987), 'A Mark 1 Geographical Analysis Machine for the automated analysis of point data sets', *International Journal of Geographical Information Systems*, 1/4: 335-358.

Openshaw, S., Craft, A.W., Charlton, M. and Birch, J.M. (1988), 'Investigation of leukaemia clusters by use of a geographical analysis machine.', *Lancet*, 1: 272-273.

Openshaw, S., Cross, A., Charlton, M. and Brunsdon, C. (1990), 'Lessons learnt from a Post Mortem of a failed GIS', *2nd National Conference and Exhibition of the AGI*, Brighton.

Openshaw, S., Turton, I., MacGill, J. and Davy, J. (1998), 'Putting the Geographical Analysis Machine on the Internet', *GISRUK'98*, Edinburgh, http://www.geog.leeds.ac.uk/staff/s.openshaw/gisruk98.doc.

Ord, J.K. and Getis, A. (1995), 'Local Spatial Autocorrelation Statistics: Distributional issues and an application', *Geographical Analysis*, 27/4: 286-306.

Page, J. (1997), 'Dial M for Mapping', Mapping Awareness, 11: 25-27.

Pease, K. (1997), 'Crime Prevention', in m. Maguire, R. Morgan and R. Reiner, eds., *The Oxford Handbook of Criminology*. 1267. Clarendon Press: Oxford.

Peuquet, D.J. (1994), 'Its about time - a conceptual-framework for the representation of temporal dynamics in Geographical Information Systems', *Annals of the Association of American Geographers*, 84/3: 441-461.

Peuquet, D.J. and Niu, D.A. (1995), 'An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data', *International Journal of Geographical Information Systems*, 9/1: 7-24.

Polvi, N., Looman, T., Humphries, C. and Pease, K. (1990), 'Repeat break-and-enter victimization: Time course and crime prevention opportunity', *Journal of Police Science and Administration*, 17/1: 8-11.

Polvi, N., Looman, T., Humphries, C. and Pease, K. (1991), 'The time course of repeat burglary victimization', *British Journal of Criminology*, 31/4: 411-414.

Pyle, D. and Deadman, D. (1994), 'Crime and unemployment in Scotland: some further results', *Scottish Journal of Political Economy*, 41/3: 314-324.

Raafat, H., Yang, Z.S. and Gauthier, D. (1994), 'Relational spatial topologies for historical geographical information', *International Journal of Geographical Information Systems*, 8/2: 163-173.

Raper, J., Rhind, D.W. and Shepherd, J.W. (1992), *Postcodes: the new geography*. London: Longman.

Ratcliffe, J.H. and McCullagh, M.J. (1998), 'Aoristic crime analysis', *International Journal of Geographical Information Science*, 12/7: 751-764.

Ratcliffe, J.H. and McCullagh, M.J. (1998a), 'Identifying repeat victimisation with GIS', *British Journal of Criminology*, 38/4: 651-662.

Ratcliffe, J.H. and McCullagh, M.J. (1998b), 'The perception of crime hot spots: A spatial study in Nottingham, UK', in N. LaVigne and J. Wartell, eds., *Crime Mapping Case Studies: Successes in the Field.* 134. Police Executive Research Forum: Washington DC.

Ratcliffe, J.H. and McCullagh, M.J. (1999), 'Burglary, Victimisation and Social Deprivation', *Crime Prevention and Community Safety*, 1/2: 37-46.

Ratcliffe, J.H. and McCullagh, M.J. (in lit.), 'Hotbeds of crime and the search for spatial accuracy', *Geographical Systems.*

Read, T. and Oldfield, D. (1995), 'Local crime analysis', *Police Research Group: Crime Detection and Prevention Series*, Paper 65: 61.

Reilly, B. and Witt, R. (1992), 'Crime and unemployment in Scotland', *Scottish Journal of Political Economy*, 39/2: 213-228.

Reiner, R. (1997), 'Policing and the police', in M. Maguire, R. Morgan and R. Reiner, eds., *The Oxford Handbook of Criminology.* 1267. Clarendon Press: Oxford.

Salinas, Police Department (1997), 'Law enforcement use of Geographic Information Systems', http://www.salinaspd.com/gis_vb.html.

Sampson, A. and Phillips, C. (1995), 'Reducing repeat racial victimisation on an East London estate', *Police Research Group: Crime Detection and Prevention Series*, Paper 67: 53.

Shaw, C.R. and McKay, H.D. (1942), *Juvenile delinquency and urban areas.* Chicago: Chicago University Press.

Sherman, L., Schmidt, J., Rogan, D., Gartin, P., Cohn, E. (1991), 'From initial deterrence to long term escalation: Short-custody arrest for poverty ghetto violence', *Criminology*, 29: 774-802.

Smith, P. (1997), 'High performance ambulances provide an intelligent response', Mapping Awareness, 11: 19-21.

Smith, S.J. (1986), *Crime, space and society.* Cambridge: Cambridge University Press.

South, N. (1997), 'Drugs: Use, Crime, and Control', in M. Maguire, R. Morgan and R. Reiner, eds., *The Oxford Handbook of Criminology.* 925-960. Clarendon Press: Oxford.

Spelman, W. (1995), 'Once bitten, then what - cross-sectional and time-course explanations of repeat victimization', *British Journal of Criminology*, 35/3: 366-383.

Tempe Police Department (1997), 'Tempe Police Department's Crime Analysis Unit'. Tempe, Arizona, http://www.tempe.gov/cau/default.htm.

Trickett, A., Osborn, D.R., Seymour, J. and Pease, K. (1992), 'What is different about high crime areas', *British Journal of Criminology*, 32/1: 81-89.

Unwin, D.J. (1996), 'GIS, spatial analysis and spatial statistics', *Progress in Human Geography*, 20/4: 540-441.

Weisburd, D. and McEwen, T. (1998), *Crime mapping and crime prevention.* New York: Criminal Justice Press.

Williamson, D.T. (1998), 'Assistant Chief Constable (Designate), Nottinghamshire Constabulary'.

Wrighton, T. (1987), 'Crime pattern analysis - an application on the Police National Computer', *Journal of the Forensic Science Society*, 27/5: 349.

Appendices

Appendix A

Table of data columns available from Nottinghamshire Police crime records.

Appendix B

Table of data columns available from Nottinghamshire Police incident (calls for service) records.

Appendix C

Standard measure of Getis and Ord G_i and G_i^* statistic.

Appendix D

MapBasic SQL script for aoristic searches.

Appendix E

Instruction pages for Trent division police perception survey.

APPENDIX A: DATA AVAILABLE FROM NOTTINGHAMSHIRE POLICE CRIME RECORDS

Column	Code	Remarks
1	Division	Police division where crime committed
2	Year	
3	CrimeNo	Unique crime reference number
4	HOmajor	Home Office major classification
5	HOminor	Home Office minor classification
6	FORCEmajor	Notts Constabulary major classification
7	FORCEminor	Notts Constabulary minor classification
8	FROMdate	
9	TOdate	
10	FROMtime	
11	TOtime	
12	Postcode	
13	ADDcode	Notts Constabulary original gazetteer code
14	STREETcode	Notts Constabulary original gazetteer code
15	DISTcode	Notts Constabulary original gazetteer code
16	GRIDcode	OS National Grid reference
17	Beat	
18	Station	
19	PREMcode	Numerical code for type of premises
20	PREMsub	PREMcode sub-code
21	MOEcode	Numerical code for method of entry
22	MOEsub	MOEcode sub-code
23	POEcode	Numerical code for point of entry
24	POEsub	POEcode sub-code
25	PROPdamaged	Value of property damaged
26	PROPstolen	Value of property stolen
27	Ethnicity	Victim details
28	Gender	"
29	Age	"
30	LookupVehicleCode	Details of any vehicle involved
31	VEHindex	Details of any vehicle involved (/cont.)
32	VEHmodel	"
33	VEHmake	"
34	VEHtype	"
35	VEHclass	"
36	VEHcolour1	"
37	VEHcolour2	"

38	ORGtype	Victim details (if an organisation)
39	ORGname	"
40	ADDbldg	Victim address details
41	ADDnumber	"
42	ADDfloorNo	"
43	ADDsubUnitNo	"
44	ADDsubUnitName	"
45	ADDsubStreet	"
46	ADDstreet	"
47	ADDdistrict	"
48	ADDcity	"
49	ADDtown	"

Source: Nottinghamshire Constabulary Crime Recording Interim System.

APPENDIX B: DATA AVAILABLE FROM NOTTINGHAMSHIRE POLICE INCIDENT RECORDS

Column	Code	Remarks
1	DIVISION_CODE	Divisional code of call origin
2	GLOBAL_INC_NO	Force unique reference number
3	REC_DATE	Date of incident
4	REC_TIME	Time of incident
5	INCIDENT_GRADE	Response grading (see chapter 3)
6	CALL_SOURCE	E.G. Ordinary phone call, 999 system etc
7	INCIDENT_TYPE	Text field describing brief nature of the incident
8	REV_INCIDENT_TYPE	Any subsequent revisions to above field
9	LOCATION	Beat code
10	GRIDREF	Full OS grid reference
11	XCOORD	Extracted Easting
12	YCOORD	Extracted Northing
13	INITIAL_RESULT	Result of the incident
14	RESULT_CODE1	Any subsequent revision to above
15	RESULT_CODE2	"
16	RESULT_CODE3	"
17	RESULT_CODE4	"
18	RESULT_CODE5	"
19	RESULT_CODE6	"
20	LONG_TIME_OF_REC_TIME	Incident date & time in format: 11/01/1997 15:38:35
21	CONCAT_DATE_TIME	Incident date & time in format: 11/01/97 15:38
22	MAJ_CAT_NO	Force incident category (see chapter 3)
23	DOW	Day of the week
24	TIME_OF_INC	Incident time in format: 1538

Source: Nottinghamshire Constabulary Command and Control System.

APPENDIX C: STANDARD MEASURE OF GETIS AND ORD'S G_I STATISTIC

n	0.90	0.95	0.99	0.999
1	1.2816	1.645	2.3238	3.0917
2	1.6323	1.9546	2.5759	3.2944
3	1.8183	2.1214	2.713	3.41
4	1.9432	2.2342	2.8067	3.49
5	2.0367	2.3189	2.8769	3.55
6	2.1107	2.3865	2.9364	3.5875
7	2.1718	2.4425	2.984	3.6286
8	2.2239	2.49	3.0225	3.675
9	2.2692	2.5319	3.0583	3.6927
10	2.3091	2.5683	3.0888	3.7248
11	2.3447	2.6015	3.1189	3.7475
12	2.3767	2.6304	3.1443	3.7655
13	2.406	2.6575	3.1708	3.7854
14	2.4329	2.6827	3.1893	3.803
15	2.4575	2.706	3.2091	3.8212
16	2.4806	2.7274	3.2282	3.8367
17	2.5018	2.7478	3.2455	3.8509
18	2.5219	2.7656	3.2621	3.8659
19	2.5413	2.7825	3.2784	3.8791
20	2.5594	2.7993	3.2931	3.8909
30	2.6964	2.9291	3.405	3.9885
40	2.7913	3.0175	3.49	4.0551
50	2.8631	3.0833	3.5488	4.1073
60	2.9218	3.14	3.5906	4.1488
100	3.0778	3.2889	3.7238	4.2659
500	3.5375	3.7134	4.1075	4.62
1000	3.7062	3.8855	4.2643	4.7667

Source: Ord and Getis (1995; p.297)

APPENDIX D: MAPBASIC SCRIPT FOR AORISTIC SQL SEARCHES

```
//--------------------------------------------------------
//                          RIGID SEARCHES
//--------------------------------------------------------
//The script for a rigid selection of just those records
//which are within the correct date criteria is simple.

SELECT * from database where
(database.FromDate  >=  search.FromDate)  AND  (database.ToDate  <=
search.ToDate)
into result1

//If the query also incorporates a search for specific
//time periods, then an additional selection is necessary
//and becomes more complicated.
//The following allows for the police officer or researcher
//to search for all incidents over a specific time period.
//
//If the time criteria are on the same day, for example
//2pm-6pm (1400-1800hrs) then a same day search is needed.

If (search.FromTime <= search.ToTime) then
SELECT * from result1 where
((result1.FromDate = result1.ToDate) AND
(result1.FromTime >= search.FromTime) AND
(result1.ToTime <= search.ToTime) AND
(result1.ToTime >= result1.FromTime))
       into FinalResult
End if

//If however the time criteria are for an overnight search,
//for example 10pm-6am (2200-0600hrs) then the script has
//to reflect this overnight component.

IF search.FromTime > search.ToTime then
SELECT * from result1 where
((result1.ToDate = result1.FromDate +1) AND
(result1.FromTime >= search.FromTime)AND
(result1.ToTime <=search.ToTime))
AND
((result1.ToDate = result1.FromDate) AND
(result1.FromTime <= search.ToTime) AND
(result1.ToTime <= search.ToTime))
AND
((result1.ToDate= result1.FromDate) AND
(result1.FromTime >= search.FromTime) AND
(result1.ToTime >= search.FromTime))
into FinalResult

//FinalResult now contains all the records from a
//rigid temporal date and time search criteria
//
//--------------------------------------------------------
//                          AORISTIC SEARCHES
//--------------------------------------------------------
```

```
//
//Script for search to include aoristic temporal matches
//Start with selecting the records with possible date matches

SELECT * from database where
((database.FromDate >= search.FromDate) AND (database.FromDate <=
search.ToDate))
OR
((database.ToDate >= search.FromDate) AND (database.ToDate <=
search.ToDate))
OR
((database.FromDate < search.FromDate) AND (database.ToDate >
search.ToDate))
into result1

//The following allows for the police officer or researcher
//to search for all incidents over, for example, the last
//month that might have occurred over a specific time period
//like night duty (10pm - 6am).

SELECT * from result1 where
((result1.FromTime >= search.FromTime) AND
(result1.FromTime <= search.ToTime))
OR
((result1.ToTime >= search.FromTime) AND
(result1.ToTime <= search.ToTime))
OR
((result1.FromTime <= search.ToTime) AND
(search.FromTime > search.ToTime))          //overnight
OR
((result1.ToTime >= search.FromTime) AND
(search.FromTime > search.ToTime))          //overnight
OR
((result1.ToTime >= search.ToTime) AND
(result1.FromDate < result1.ToDate))
OR
((result1.FromTime <= search.FromTime) AND
(result1.ToDate > result1.FromDate))
OR
(result1.ToDate >= result1.FromDate+ 2)
into FinalResult

//FinalResult now contains all the records from a
// aoristic temporal date and time search criteria
```

APPENDIX E: INSTRUCTION PAGES FOR TRENT DIVISION POLICE PERCEPTION SURVEY

The following two pages are the instruction sheets given to Trent division officers on how to complete the hotspot perception study described in Chapter 9. The two pages instructions were followed by four maps of the officer's subdivision (either Clifton, Meadows or West Bridgford). The four target crime types were: residential burglary, non-residential burglary, assaults and disorder, and motor vehicle crime.

The formatting has been adjusted slightly for the different page layout of this thesis.

Crime Hotspot Survey 1998

This survey aims to examine the knowledge of people in the police service regarding the pattern of crime in their work area. As we all know, crime is not distributed evenly over an area and Clifton is no different. There are areas of Clifton which experience more crime than others and these crime concentrations, or 'hotspots', are the areas which this survey is interested in. We would like to know where you think the crime hotspots have been over the last six months.

This sheet will tell you how to complete the survey and the following sheets show maps on which we would like you to mark where you think the hotspots are. Don't worry - it is not a test, and you can not get it wrong! The final sheet has a few questions we would like you to answer about yourself, and the whole survey should only take about 6 or 7 minutes to complete. Thanks for helping.

The survey

A hotspot is an area where crime appears to be more highly concentrated than other areas. On the following pages are a number of identical maps showing Clifton. **At the top of each page is the type of crime that the page refers to.** The types of activity this survey is interested in follow:

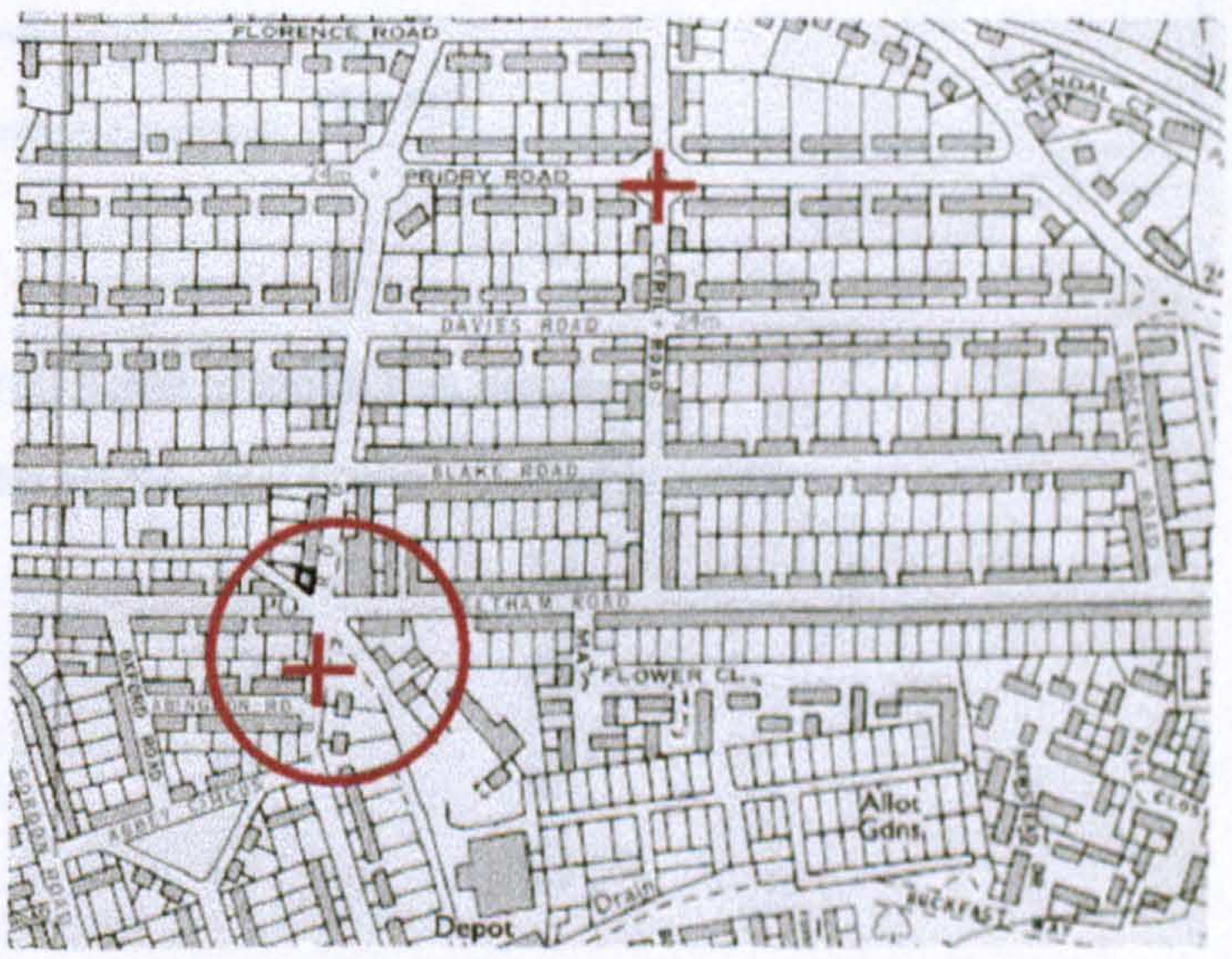
- Residential burglary (dwelling)
- Non-residential burglary (commercial premises, schools, etc.)
- Motor vehicle crime
- Street disturbance/rowdyism

When thinking about crimes like burglary, include attempts and crimes which resulted in arrests. Motor vehicle crime includes TWOC, theft from, and all attempts.

Please do not work on this with anyone else. We want to know your ideas - not a team effort!

How to complete the survey

1. On the first sheet, think about where in the station area you think the highest concentration ('hotspot') of residential burglaries might be, based on your experience. Place a clearly marked cross on the map at the centre of the hotspot area. See the example on the right.
2. Now draw a circle or other shape around your cross to show the extent of that hotspot. For example if you think it only extends a little bit around that area, draw a small circle. If it takes in a number of the surrounding streets, make the shape larger.
3. If there is more than one hotspot in the station area, show the other area(s) with crosses as well - **but do not mark them with a circle**. This will enable us to see which is your most important 'hotspot'. Also do not try to show every location where a crime has taken place! The survey is only interested in the areas of concentration. See example below.



4. Repeat this method for each of the maps.
5. Finally, please fill in the final sheet asking a few questions about yourself and this survey.

Note: the shape around the hotspot does not need to be a circle.

Once again - thanks for your help!

Please take the time to fill in these few questions about yourself and this survey as it will help us to more understand the information you have given us.

1. How long have you worked for the police (not just Notts police but anywhere)?____ YEARS

2. How long have you worked at your current station? ____________ YEARS

3. Are you police officer or civilian employee? ☐ POLICE ☐ CIVILIAN

4. **If you are a police officer,** which category fits you best?

☐ Uniform operational officer

☐ CID

☐ PBO / Community officer

☐ Station staff

☐ Other. Describe ____________________________

5. What is your job? (i.e. schools liaison, detective etc.)

____

6. What rank are you? ______________

7. **If you are a civilian employee,** what is your job?

____

8. How long have you done this job? __________ YEARS

About this survey...

9. How confident are you about the answers you have given in this survey?

	Burglary dwelling	Burglary non-res	MV crime	Disturbance
Very confident	☐	☐	☐	☐
Fairly confident	☐	☐	☐	☐
Not too confident	☐	☐	☐	☐
Very unconfident	☐	☐	☐	☐

Please fill in any further comments you have about the survey...

____

____

____

Thanks for your help. If no-one is available to collect your survey questionnaire at present then please either give it to your supervisor to return to West Bridgford or send it to the "Crime Hotspot Survey" at West Bridgford station in the internal post.

NO CD

ATTACHED

PLEASE APPLY

TO

UNIVERSITY